

## FIGURE 1

ACTGCACCTCGGTTCTTGGATTGAATTCCCCGGGGATCCTCTAGAGATCCCTCGACCTCGA  
CCCACGCGTCCGGGCGGACAGCACGGCCGAGGACCTGGAGCTCCGGCTGCGTCTTCCCG  
CAGCGCTACCCGCCATGCGCCTGCCGCGCCGGGCCGCGCTGGGGCTCCTGCCGCTTCTGCTG  
CTGCTGCCGCCCCGCGCCCGAGGCCGCCAAGAAGCCGACGCCCTGCCACCGGTGCCGGGGGCT  
GGTGGACAAGTTTAACCAGGGGATGGTGGACACCGCAAAGAAGAACTTTGGCGGCGGGAACA  
CGGCTTGGGAGGAAAAGACGCTGTCCAAGTACGAGTCCAGCGAGATTGCCTGCTGGAGATC  
CTGGAGGGGCTGTGCGAGAGCAGCGACTTCGAATGCAATCAGATGCTAGAGGCGCAGGAGGA  
GCACCTGGAGGCCTGGTGGCTGCAGCTGAAGAGCGAATATCCTGACTTATTCGAGTGGTTTT  
GTGTGAAGACACTGAAAGTGTGCTGCTCTCCAGGAACCTACGGTCCCGACTGTCTCGCATGC  
CAGGGCGGATCCCAGAGGCCCTGCAGCGGGAATGGCCACTGCAGCGGAGATGGGAGCAGACA  
GGGCGACGGGTCTGCCGGTGGCCACATGGGGTACCAGGGCCCCGCTGTGCACTGACTGCATGG  
ACGGCTACTTCAGCTCGCTCCGGAACGAGACCCACAGCATCTGCACAGCCTGTGACGAGTCC  
TGCAAGACGTGCTCGGGCCTGACCAACAGAGACTGCGGCGAGTGTGAAGTGGGCTGGGTGCT  
GGACGAGGGCGCCTGTGTGGATGTGGACGAGTGTGCGGCCGAGCCGCCTCCCTGCAGCGCTG  
CGCAGTTCTGTAAGAACGCCAACGGCTCCTACACGTGCGAAGAGTGTGACTCCAGCTGTGTG  
GGCTGCACAGGGGAAGGCCCAGGAACTGTAAAGAGTGTATCTCTGGCTACGCGAGGGAGCA  
CGGACAGTGTGCAGATGTGGACGAGTGTCTACTAGCAGAAAAAACCTGTGTGAGGAAAAACG  
AAAAGTCTACAATACTCCAGGGAGCTACGTCTGTGTGTGTCTGACGGCTTCGAAGAAACG  
GAAGATGCCTGTGTGCCGCCGCGCAGAGGCTGAAGCCACAGAAGGAGAAAGCCCGACACAGCT  
GCCCTCCCGCGAAGACCTGTAAATGTGCGGACTTACCCTTTAAATTATTCAGAAGGATGTCC  
CGTGGAATGTGGCCCTGAGGATGCCGTCTCCTGCAGTGGACAGCGGCGGGGAGAGGCTGC  
CTGCTCTCTAACGGTTGATTCTCATTTGTCCCTTAAACAGCTGCATTTCTTGGTTGTTCTTA  
AACAGACTTGTATATTTTGATACAGTTCTTTGTAATAAAATTGACCATTGTAGGTAATCAGG  
AGGAAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCGACCTGCAGAAGC  
TTGGCCGCCATGGCCCAACTTGTTTATTGCAGCTTATAATGGTTACAAATAAAGCAATAGCA  
TCACAAATTCACAAATAAAGCATTTTTTTTCACTGCATTCTAGTTGTGGTTTGTCCAACTC  
ATCAATGTATCTTATCATGTCTGGATCGGGAATTAATTCCGGCGCAGCACCATGGCCTGAAAT  
AACCTCTGAAAGAGGAACTTGGTTAGGTACCTTCTGAGGCGGAAAGAACCAGCTGTGGAATG  
TGTGTCAGTTAGGGTGTGGAAAGTCCCCAGGCTCCCCAGCAGGCAGAAGTATGCAAGCATGC  
ATCTCAATTAGTCAGCAACCCAGTTTTT

CCCTCGACCTCGA

## FIGURE 2

><subunit 1 of 1, 353 aa, 0 stop

><MW: 38192, pI: 4.53, NX(S/T): 2

MRLPRRAALGLLPLLLLLLPPAPEAAKKPTPCHRCRGLVDKFNQGMVDTAKKNFGGGNTAWEKTLISKYESSSEIRL  
LEILEGLCESSDFECNQMLEAQEEHLEAWWLQLKSEYPDLFEWFCVKTLKVCCSPGTYGPDCLACQGSQRPCSG  
NGHCSGDGSRQGDGSCRCHMGYQGPLCTDCMDGYFSSLRNETHSICTACDESKTCSGLTNRDCGECEVGWVLDE  
GACVDVDECAAEPPPCSAAQFCKNANGSYTCEECDSSCVGCTGEGPGNCKECISGYAREHGQCADVDECSLAEKT  
CVRKNENCYNTPGSYVCVCPDGFEEETEDACVPPAEAEATEGESPTQLPSREDL

### **Signal peptide:**

amino acids 1-24

### **N-glycosylation sites.**

amino acids 190-194 and 251-255

### **Glycosaminoglycan attachment sites.**

amino acids 149-153 and 155-159

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 26-30

### **Casein kinase II phosphorylation sites.**

amino acids 58-62, 66-70, 86-90, 197-201, 210-214, 255-259, 295-299, 339-343  
and 349-353

### **Tyrosine kinase phosphorylation site.**

amino acids 303-310

### **N-myristoylation sites.**

amino acids 44-50, 54-60, 55-61, 81-87, 150-156, 158-164, 164-170, 252-258 and  
313-319

### **Aspartic acid and asparagine hydroxylation site.**

amino acids 308-320

### **EGF-like domain cysteine pattern signature.**

amino acids 166-178

### **Leucine zipper pattern.**

amino acids 94-116

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

### FIGURE 3

CAGGTCCAACCTGCACCTCGGTTCTATCGATTGAATTCCCCGGGGATCCTCTAGAGATCCCTC  
GACCTCGACCCACGCGTCCGCCAGGCCGGGAGGCGACGCGCCAGCCGTCTAAACGGGAACA  
GCCCTGGCTGAGGGAGCTGCAGCGCAGCAGAGTATCTGACGGCGCCAGGTTGCGTAGGTGCG  
GCACGAGGAGTTTTCCCGGCAGCGAGGAGGTCTTGAGCAGCATGGCCCCGGAGGAGCGCCTTC  
CCTGCCGCCGCGCTCTGGCTCTGGAGCATCCTCCTGTGCCTGCTGGCACTGCGGGCGGAGGC  
CGGGCCGCCGAGGAGGAGAGCCTGTACCTATGGATCGATGCTCACCAGGCAAGAGTACTCA  
TAGGATTTGAAGAAGATATCCTGATTGTTTTAGAGGGGAAAATGGCACCTTTTACACATGAT  
TTCAGAAAAGCGCAACAGAGAATGCCAGCTATTCTGTCAATATCCATTCCATGAATTTTAC  
CTGGCAAGCTGCAGGGCAGGCAGAATACTTCTATGAATTCCTGTCCTTGCGCTCCCTGGATA  
AAGGCATCATGGCAGATCCAACCGTCAATGTCCCTCTGCTGGGAACAGTGCCTCACAAGGCA  
TCAGTTGTTCAAGTTGGTTTTCCCATGTCTTGGAACACAGGATGGGGTGGCAGCATTGAAAGT  
GGATGTGATTGTTATGAATTCTGAAGGCAACACCATTCTCCAAACACCTCAAAATGCTATCT  
TCTTTAAACATGTCAACAAGCTGAGTGCCAGGCGGGTGGCGAAATGGAGGCTTTTGTAAT  
GAAAGACGCATCTGCGAGTGTCCTGATGGGTTCCACGGACCTCACTGTGAGAAAGCCCTTTG  
TACCCACGATGTATGAATGGTGGACTTTGTGTGACTCCTGGTTTCTGCATCTGCCACCTG  
GATTCTATGGAGTGAACGTGACAAAGCAAACCTGCTCAACCACCTGCTTTAATGGAGGGACC  
TGTTTCTACCCTGGAAAATGTATTTGCCCTCCAGGACTAGAGGGAGAGCAGTGTGAAATCAG  
CAAATGCCCAACCCCTGTGCAATGGAGGTAAATGCATTGGTAAAAGCAAATGTAAGTGTT  
CCAAAGGTTACCAGGGAGACCTCTGTTCAAAGCCTGTCTGCGAGCCTGGCTGTGGTGCACAT  
GGAACCTGCCATGAACCCAACAAATGCCAATGTCAAGAAGGTTGGCATGGAAGACACTGCAA  
TAAAAGGTACGAAGCCAGCCTCATACATGCCCTGAGGCCAGCAGGCGCCAGCTCAGGCAGC  
ACACGCCTTCACTTAAAAAGGCCGAGGAGCGGCGGGATCCACCTGAATCCAATTACATCTGG  
TGAACCTCCGACATCTGAAACGTTTTAAGTTACACCAAGTTCATAGCCTTTGTAAACCTTTCA  
TGTGTTGAATGTTCAAATAATGTTTATTACACTTAAGAATACTGGCCTGAATTTTATTAGCT  
TCATTATAAATCACTGAGCTGATATTTACTCTTCCTTTTAAGTTTTCTAAGTACGTCTGTAG  
CATGATGGTATAGATTTTCTTGTTTTCAGTGCTTTGGGACAGATTTTATATTATGTCAATTGA  
TCAGGTTAAAATTTTTCAGTGTGTAGTTGGCAGATATTTTCAAATTAACAATGCATTTATGGT  
GTCTGGGGGCAGGGGAACATCAGAAAGGTTAAATTTGGGCAAAAATGCGTAAGTCACAAGAAT  
TTGGATGGTGCAGTTAATGTTGAAGTTACAGCATTTCAGATTTTATTGTCAGATATTTAGAT  
GTTTGTACATTTTAAAAAATTGCTCTTAATTTTAACTCTCAATACAATATATTTTGACC  
TTACCATTATTCCAGAGATTCAGTATTAAAAAATAAATTAACCTGTGGTAGTGGCATT  
AAACAATATAATATATTCTAAACACAATGAAATAGGGAATATAATGTATGAACTTTTGCAT  
TGGCTTGAAGCAATATAATATATTGTAAACAAAACACAGCTCTTACCTAATAAACATTTTAT  
ACTGTTTGTATGTATAAAATAAAGGTGCTGCTTTAGTTTTTTGGAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAGAGTCGACCTGCAGAAGCTTGGC  
CGCCATGGCCCACTTGTTTATTGCAGCTTATAATG

## FIGURE 4

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA33094

><subunit 1 of 1, 379 aa, 0 stop

><MW: 41528, pI: 7.97, NX(S/T): 2

MARRSAFPAAALWLWSILLCLLALRAEAGPPQEESELYLWIDAHQARVLIGFEEDILIVSEGK  
MAPFTHDFRKAQQRMPAIPVNIHSMNFTWQAAGQAEYFYEFSLRSLDKGIMADPTVNVPLL  
GTVPHKASVVQVGFPCLGKQDGVAAFEVDVIVMNSEGNTILQTPQNAIFFKTCQQAECPPGGC  
RNGGFCNERRICECPDGFHGHCEKALCTPRCMNGGLCVTPGFCICPPGFYGVNCDKANCST  
TCFNGGTCTFYPGKICPPGLEGEQCEISKCPQPCRNGGKICIGKSKCKCSKGYQGDLCCKPVC  
EPGCGAHGTCHEPNKCQCQEGWHGRHCNKRYEASLIHALRPAGAQLRQHTPSLKKAERRDP  
PESNYIW

### **Signal peptide:**

amino acids 1-28

### **N-glycosylation site.**

amino acids 88-92, 245-249

### **Casein kinase II phosphorylation site.**

amino acids 319-323

### **Tyrosine kinase phosphorylation site.**

amino acids 370-378

### **N-myristoylation sites.**

amino acids 184-190, 185-191, 189-195, 315-321

### **ATP/GTP-binding site motif A (P-loop).**

amino acids 285-293

### **EGF-like domain cysteine pattern signature.**

amino acids 198-210, 230-242, 262-274, 294-306, 326-338

protein database



## FIGURE 5

CGGACGCGTGGGCGTCCGGCGGTTCGACAGAGCCAGGAGGCGGAGGCGCGCGGGCCAGCCTGGG  
CCCCAGCCCACACCTTCACCAGGGCCCCAGGAGCCACCATGTGGCGATGTCCACTGGGGCTAC  
TGCTGTTGCTGCCGCTGGCTGGCCACTTGGCTCTGGGTGCCCAGCAGGGTCTGTTGGGCGCCGG  
GAGCTAGCACCGGGTCTGCACCTGCGGGGCATCCGGGACGCGGGAGGCCGGTACTGCCAGGA  
GCAGGACCTGTGCTGCCGCGGCCGTGCCGACGACTGTGCCCTGCCCTACCTGGGCGCCATCT  
GTTACTGTGACCTCTTCTGCAACCGCACGGTCTCCGACTGCTGCCCTGACTTCTGGGACTTC  
TGCCTCGGCGTGCCACCCCCCTTTTCCCCCGATCCAAGGATGTATGCATGGAGGTCTGATCTA  
TCCAGTCTTGGGAACGTACTGGGACAACGTGAACCGTTGCACCTGCCAGGAGAACAGGCAGT  
GGCATGGTGGATCCAGACATGATCAAAGCCATCAACCAGGGCAACTATGGCTGGCAGGCTGG  
GAACCACAGCGCCTTCTGGGGCATGACCCTGGATTGAGGGCATTTCGCTACCGCCTGGGCACCA  
TCCGCCCATCTTCTCGGTCTATGAACATGCATGAAATTTATACAGTGCTGAACCCAGGGGAG  
GTGCTTCCCACAGCCTTCGAGGCCTCTGAGAAGTGGCCCAACCTGATTCATGAGCCTCTTGA  
CCAAGGCAACTGTGCAGGCTCCTGGGCCTTCTCCACAGCAGCTGTGGCATCCGATCGTGTCT  
CAATCCATTCTCTGGGACACATGACGCCTGTCTGTGCGCCCAGAACCTGCTGTCTTGTGAC  
ACCCACCAGCAGCAGGGCTGCCGCGGTGGGCGTCTCGATGGTGCCTGGTGGTTCCTGCGTCG  
CCGAGGGGTGGTGTCTGACCACTGCTACCCCTTCTCGGGCCGTGAACGAGACGAGGCTGGCC  
CTGCGCCCCCTGTATGATGCACAGCCGAGCCATGGGTGCGGGCAAGCGCCAGGCCACTGCC  
CACTGCCCCAACAGCTATGTTAATAACAATGACATCTACCAGGTCACTCCTGTCTACCGCCT  
CGGCTCCAACGACAAGGAGATCATGAAGGAGCTGATGGAGAATGGCCCTGTCCAAGCCCTCA  
TGGAGGTGCATGAGGACTTCTTCTTATAACAAGGAGGCATCTACAGCCACACGCCAGTGAGC  
CTTGGGAGGCCAGAGAGATAACCGCCGGCATGGGACCCACTCAGTCAAGATCACAGGATGGGG  
AGAGGAGACGCTGCCAGATGGAAGGACGCTCAAATACTGGACTGCGGCCAACTCCTGGGGCC  
CAGCCTGGGGCGAGAGGGGGCCACTTCCGCATCGTGCGCGGCGTCAATGAGTGCGACATCGAG  
AGCTTCGTGCTGGGCGTCTGGGGCCGCGTGGGCATGGAGGACATGGGTCTACTGAGGCTG  
CGGGCACCACGCGGGGTCCGGCCTGGGATCCAGGCTAAGGGCCGGCGGAAGAGGGCCCCAATG  
GGGCGGTGACCCAGCCTCGCCCGACAGAGCCCGGGGCGCAGGCGGGCGCCAGGGCGCTAAT  
CCCGGCGCGGGTTCCGCTGACGACGCGCCCCGCTGGGAGCCGCGGGCAGGCGAGACTGGCG  
GAGCCCCCAGACCTCCAGTGGGGACGGGGCAGGGCCTGGCCTGGGAAGAGCACAGCTGCAG  
ATCCCAGGCCTCTGGCGCCCCCACTCAAGACTACCAAAGCCAGGACACCTCAAGTCTCCAGC  
CCCAATACCCACCCCAATCCCGTATTCTTTTTTTTTTTTTTTTTTTAGACAGGGTCTTGCTCCG  
TTGCCCAGGTTGGAGTGAGTGAGTGGCCATCAGGGCTCACTGTAACCTCCGACTCCTGGGTTCA  
AGTGACCCTCCCACCTCAGCCTCTCAAGTAGCTGGGACTACAGGTGCACCACCACACCTGGC  
TAATTTTTGTATTTTTTTGTAAAGAGGGGGGTCTCACTGTGTTGCCAGGCTGGTTTCGAACT  
CCTGGGCTCAAGCGGTCCACCTGCCTCCGCCTCCCAAAGTGCTGGGATTGCAGGCATGAGCC  
ACTGCACCCAGCCCTGTATTCTTATTCTTCAGATATTTATTTTTCTTTTCACTGTTTTAAAA  
TAAAACCAAAGTATTGATAAAAAAAAAA

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## **FIGURE 6**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA33223

><subunit 1 of 1, 164 aa, 1 stop

><MW: 18359, pI: 7.45, NX(S/T): 1

MWRCPLGLLLLLLPLAGHLALGAQQGRGRRELAPGLHLRGIRDAGGRYCQEQLCCRGRADDC  
ALPYLGAICYCDLFCNRTVSDCCPDFWDFCLGVPPFPFPIQGCMHGGRIYPVLGTYWDNCNR  
CTCQENRQWHGGSRHDQSHQPGQLWLAGWEPQRLLGHDPG

**N-glycosylation site.**

amino acids 78-82, 161-165

**Casein kinase II phosphorylation site.**

amino acids 80-84, 117-121, 126-130, 169-173, 205-209, 296-300,  
411-415

**N-myristoylation site.**

amino acids 21-27, 39-45, 44-50, 104-110, 160-164, 224-230,  
269-275, 378-384, 442-448

**Amidation site.**

amino acids 26-30, 318-322

**Eukaryotic thiol (cysteine) proteases histidine active site.**

amino acids 398-409

protein 2000

## FIGURE 7

AGGCTCCTTGGCCCTTTTTCCACAGCAAGCTTNTGCNATCCCGATTTCGTTGTCTCAAATCCA  
ATTCTCTTGGGACACATNACGCCTGTCCTTTNGCCCCAGAACCTGCTGTCTTGTACACCCAC  
CAGCAGCAGGGCTGCCGCGNTGGGCGTCTCGATGGTGCCTGGTGGTTCCTGCGTCGCCGAGG  
GNTGGTGTCTGACCACTGCTACCCCTTCTCGGGCCGTGAACGAGACGAGGCTGGCCCTGCGC  
CCCCCTGTATGATGCACAGCCGAGCCATGGGTGCGGGCAAGCGCCAGGCCACTGCCCCTGCGC  
CCCAACAGCTATGTTAATAACAATGACATCTACCAGGTCCTGTCTACCGCCTCGGCTC  
CAACGACAAGGAGATCATGAAGGAGCTGATGGAGAATGGCCCTGTCCAAGCCCTCATGGAGG  
TGCATGAGGACTTCTTCCTATAACAAGGAGGCATCTACAGCCACACGCCAGTGAGCCTTGGG  
AGGCCAGAGAGATACCGCCGGCATGGGACCCACTCAG

FIG. 29. 61130600

## **FIGURE 8**

GCTGCTTGCCCTGTTGATGGCAGGCTTGGCCCTGCAGCCAGGCACTGCCCTGCTGTGCTACT  
CCTGCAAAGCCCAGGTGAGCAACGAGGACTGCCTGCAGGTGGAGAACTGCACCCAGCTGGGG  
GAGCAGTGCTGGACCGCGCGCATCCGCGCAGTTGGCCTCCTGACCGTCATCAGCAAAGGCTG  
CAGCTTGAACTGCGTGGATGACTCACAGGACTACTACGTGGGCAAGAAGAACATCACGTGCT  
GTGACACCGACTTGTGCAACGCCAGCGGGGCCCATGCCCTGCAGCCGGCTGCCGCCATCCTT  
GCGCTGCTCCCTGCACTCGGCCTGCTGCTCTGGGGACCCGGCCAGCTATAGGCTCTGGGGGG  
CCCCGCTGCAGCCCACACTGGGTGTGGTGCCCCAGGCCTCTGTGCCACTCCTCACAGACCTG  
GCCCAGTGGGAGCCTGTCCTGGTTCCTGAGGCACATCCTAACGCAAGTCTGACCATGTATGT  
CTGCACCCCTGTCCCCCACCCCTGACCCCTCCCATGGCCCTCTCCAGGACTCCCACCCGGCAGA  
TCAGCTCTAGTGACACAGATCCGCCTGCAGATGGCCCCTCCAACCCTCTCTGCTGCTGTTTC  
CATGGCCCAGCATTTCTCCACCCCTTAACCCTGTGCTCAGGCACCTCTTCCCCCAGGAAGCCTT  
CCCTGCCCCACCCCATCTATGACTTGAGCCAGGTCTGGTCCGTGGTGTCCCCCGCACCCAGCA  
GGGGACAGGCACTCAGGAGGGCCCCAGTTAAAGGCTGAGATGAAGTGGACTGAGTAGAACTGGA  
GGACAAGAGTCGACGTGAGTTCCTGGGAGTCTCCAGAGATGGGGCCTGGAGGCCTGGAGGAA  
GGGGCCAGGCCTCACATTCGTGGGGCTCCCTGAATGGCAGCCTGAGCACAGCGTAGGCCCTT  
AATAAACACCTGTTGGATAAGCCAAAAAA

## **FIGURE 9**

MTHRTTTWARRTSRAVTPTCATPAGPMPCSRLLPSSLRCSLHSACCSGDPASYRLWGAPLQPT  
LGVVPQASVPLLTDLAQWEPVLVPEAHPNASLTMVCTPVPHDPDPPMALSRTPTRQISSSDT  
DPPADGPSNPLCCCCFHGPAFSTLNPVLRHLFPQEAFPAHPIYDLSQVWSVVSPAPSRGQALRRAQ

### **Signal peptide:**

amino acids 1-47

### **N-glycosylation site.**

amino acids 31-35, 74-78, 84-88

### **Casein kinase II phosphorylation site.**

amino acids 22-26, 76-80

### **N-myristoylation site.**

amino acids 56-60

### **Amidation site.**

amino acids 70-74

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[illegible][illegible]

## **FIGURE 11**

MGAARLLPNLTLCLQLLILCCQTQYVRDQGAMTDQLSRRQIREYQLYSRTSGKHVQVTGRRISATAEDGNKFAKLIVETDTFGSRVRIKGAESEKYICMNKRGKLGKPSGKSKDCVFTEIVLENNYTAFQONARHEGWFMATFTRQGRPRQASRSRQONQREAHFIKRLYQGQLPFPNHAEKQKQFEFVGSAPTRRTKRTRRPQPLT

### **Signal peptide:**

amino acids 1-22

### **N-glycosylation site.**

amino acids 9-13, 126-130

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 60-64

### **Casein kinase II phosphorylation site.**

amino acids 65-69

### **Tyrosine kinase phosphorylation site.**

amino acids 39-48, 89-97

### **N-myristoylation site.**

amino acids 69-75, 188-194

### **Amidation site.**

amino acids 58-62

### **HBGF/FGF family signature.**

amino acids 103-128

## FIGURE 12

ACTTGCCATCACCTGTTGCCAGTGTGGAAAAATTCTCCCTGTTGAATTTTTTGCACATGGAG  
GACAGCAGCAAAGAGGGCAACACAGGCTGATAAGACCAGAGACAGCAGGGAGATTATTTTAC  
CATACGCCCTCAGGACGTTCCCTCTAGCTGGAGTTCTGGACTTCAACAGAACCCCATCCAGT  
CATTTTGATTTTGCTGTTTATTTTTTTTTTTCTTTTCTTTTCCCACCACATTGTATTTTAT  
TTCCGTACTTCAGAAATGGGCCTACAGACCACAAAGTGGCCCAGCCATGGGGCTTTTTTCT  
GAAGTCTTGGCTTATCATTTCCCTGGGGCTCTACTCACAGGTGTCCAAACTCCTGGCCTGCC  
CTAGTGTGTGCCGCTGCGACAGGAACCTTTGTCTACTGTAATGAGCGAAGCTTGACCTCAGTG  
CCTCTTGGGATCCCGGAGGGCGTAACCGTACTCTACCTCCACAACAACCAAATTAATAATGC  
TGGATTTCCTGCAGAACTGCACAATGTACAGTCGGTGCACACGGTCTACCTGTATGGCAACC  
AACTGGACGAATTCCCCATGAACCTTCCCAAGAATGTCAGAGTTCTCCATTTGCAGGAAAAC  
AATATTAGACCATTTACGGGCTGCTCTTGCCCAGCTCTTGAAGCTTGAAGAGCTGCACCT  
GGATGACAACCTCCATATCCACAGTGGGGGTGGAAGACGGGGCCTTCCGGGAGGGCTATTAGCC  
TCAAATTGTTGTTTTTGTCTAAGAATCACCTGAGCAGTGTGCCTGTTGGGCTTCTGTGGAC  
TTGCAAGAGCTGAGAGTGGATGAAAATCGAATTGCTGTATATCCGACATGGCCTTCCAGAA  
TCTCACGAGCTTGGAGCGTCTTATTGTGGACGGGAACCTCCTGACCAACAAGGGTATCGCCG  
AGGGCACCTTCAGCCATCTCACCAGCTCAAGGAATTTTCAATTGTACGTAATTCGCTGTCC  
CACCTCCTCCCGATCTCCAGGTACGCATCTGATCAGGCTCTATTTGCAGGACAACCAGAT  
AAACCACATTCTTTGACAGCCTTCTCAAATCTGCGTAAGCTGGAACGGCTGGATATATCCA  
ACAACCAACTGCGGATGCTGACTCAAGGGGTTTTTGATAATCTCTCCAACCTGAAGCAGCTC  
ACTGCTCGGAATAACCTTGGTTTTGTGACTGCAGTATTAAATGGGTCACAGAATGGCTCAA  
ATATATCCCTTCATCTCTCAACGTGCGGGGTTTCATGTGCCAAGGTCCTGAACAAGTCCGGG  
GGATGGCCGTCAGGGAATTAATATGAATCTTTTGTCTGTCCCACCACGACCCCCGGCCTG  
CCTCTCTTACCCCCAGCCCCAAGTACAGCTTCTCCGACCACTCAGCCTCCCACCCTCTCTAT  
TCCAAACCCTAGCAGAAGCTACACGCCTCCAACCTCCTACCACATCGAAACTTCCCACGATTC  
CTGACTGGGATGGCAGAGAAAGAGTGACCCACCTATTTCTGAACGGATCCAGCTCTCTATC  
CATTTTGTGAATGATACTTCCATTCAAGTCAGCTGGCTCTCTCTCTTACCCTGATGGCATA  
CAAACCTACATGGGTGAAAATGGGCCACAGTTTAGTAGGGGGCATCGTTTACAGGAGCGCATAG  
TCAGCGGTGAGAAGCAACACCTGAGCCTGGTTAACTTAGAGCCCCGATCCACCTATCGGATT  
TGTTTAGTGCCACTGGATGCTTTTAACTACCGCGCGGTAGAAGACACCATTTGTTCAGAGGC  
CACCACCCATGCCTCCTATCTGAACAACGGCAGCAACACAGCGTCCAGCCATGAGCAGACGA  
CGTCCCACAGCATGGGCTCCCCCTTTCTGCTGGCGGGCTTGATCGGGGGCGCGGTGATATTT  
GTGCTGGTGGTCTTGCTCAGCGTCTTTTGTGCTGGCATATGCACAAAAAGGGGCGCTACACCTC  
CCAGAAGTGGAATAACAACCGGGGCGGGCGGAAAGATGATTATTGCGAGGCAGGCACCAAGA  
AGGACAACCTCCATCCTGGAGATGACAGAAACCAGTTTTTCAGATCGTCTCCTTAAATAACGAT  
CAACTCCTTAAAGGAGATTTTCAGACTGCAGCCCATTTACACCCCAAATGGGGGCATTAATTA  
CACAGACTGCCATATCCCCAACAACATGCGATACTGCAACAGCAGCGTGCCAGACCTGGAGC  
ACTGCCATACGTGACAGCCAGAGGCCAGCGTTATCAAGGCGGACAATTAGACTCTTGAGAA  
CACACTCGTGTGTGCACATAAAGACACGCAGATTACATTTGATAAATGTTACACAGATGCAT  
TTGTGCATTTGAATACTCTGTAATTTATACGGTGTACTATATAATGGGATTTAAAAAAGTG  
CTATCTTTTCTATTTCAAGTTAATTACAAACAGTTTTTGTAACCTTTTGCTTTTTTAAATCTT



## **FIGURE 13**

MGLQTTKWPSHGAFFLKSWLIISLGLYSQVSKLLACPSVCRCDRNFVYCNERSLTSVPLGIP  
EGVTVLYLHNNQINNAGFPAELHNVQSVHTVYLYGNQLDEFPMNLPKNVRVLHLQENNIQTI  
SRAALAQLLKLEELHLDNDSISTVGVEDGAFREAI SLKLLFLSKNHLSSVPVGLPVDLQELR  
VDENRIAVISDMAFQNLTSLERLIVDGNLLTNKGIAEGTFSHLTKLKEFSIVRNSLSHPPPD  
LPGTHLIRLYLQDNQINHIPLTAFSNLRKLERLDISNNQLRMLTQGVFDNLSNLKQLTARNN  
PWFCDCSIKWVTEWLKYIPSSLNVRGFMCGPEQVRGMVRELNMNLLSCPTTTPGLPLFTP  
APSTASPTTQPPTLSIPNPSRSYTPPTPTTSKLPTIPDWDGRERVTPPISERIQLSIHFVND  
TSIQVSWLSLFTVMAYKLTWVKMGHSLVGGIVQERIVSGEKQHLSLVNLEPRSTYRICLVPL  
DAFNRYRAVEDTICSEATTHASYLNNGSNTASSHEQTTSHSMGSPFLLAGLIGGAVIFVLVVL  
LSVFCWHMHKKGRYTSQKWYNRGRRKDDYCEAGTKKDNSILEMTETSFQIVSLNNDQLLKG  
DFRLQPIYTPNGGINYTDCHIPNNMRYCNSSVPDLEHCHT

**Signal peptide:**

amino acids 1-42

**Transmembrane domain:**

amino acids 542-561

**N-glycosylation site.**

amino acids 202-206, 298-302, 433-437, 521-525, 635-639, 649-653

**Casein kinase II phosphorylation site.**

amino acids 204-208, 407-411, 527-531, 593-597, 598-602, 651-655

**Tyrosine kinase phosphorylation site.**

amino acids 319-328

**N-myristoylation site.**

amino acids 2-8, 60-66, 149-155, 213-219, 220-226, 294-300,  
522-528, 545-551, 633-639

**Amidation site.**

amino acids 581-585

**Leucine zipper pattern.**

amino acids 164-186

**Phospholipase A2 aspartic acid active site.**

amino acids 39-50

# FIGURE 14

ACTTGGAGCAAGCGGCGGCGGCGGAGACAGAGGCAGAGGCAGAAAGCTGGGGCTCCGTCCCTCGCCTCCCACGAGCG  
ATCCCCGAGGAGAGCCGCGGCCCTCGGCGAGGCGAAGAGGCCGACGAGGAAGACCCGGGTGGCTGCGCCCCCTGCC  
TCGCTTCCAGGCGCCGCGGCTGCAGCCTTGCCCCCTTGCTCGCCTTGAAAATGGAAAAGATGCTCGCAGGCT  
GCTTTCTGCTGATCCTCGGACAGATCGTCCTCCTCCCTGCCGAGGCCAGGGAGCGGTACGTGGGAGGTCCATCT  
CTAGGGGCGAGACACGCTCGGACCCACCCGAGACGGCCCTTCTGGAGAGTTCTGTGAGAAACAAGCGGGCAGACC  
TGGTTTTTCATCATTTGACAGCTCTCGCAGTGTCAACACCCATGACTATGCAAAGGTCAAGGAGTTTCATCGTGGACA  
TCTTGCAATTCTTGGACATTGGTCCTGATGTACCCGAGTGGGCCTGCTCCAATATGGCAGCACTGTCAAGAATG  
AGTTCTCCCTCAAGACCTTCAAGAGGAAGTCCGAGGTGGAGCGTGTGTCAAGAGGATGCGGCATCTGTCCACGG  
GCACCATGACTGGGCTGGCCATCCAGTATGCCCTGAACATCGCATTCTCAGAAGCAGAGGGGGCCCGGCCCTGA  
GGGAGAATGTGCCACGGGTATAATGATCGTGCAGATGGGAGACCTCAGGACTCCGTGGCCGAGGTGGCTGCTA  
AGGCACGGGACACGGGCATCCTAATCTTTGGCATTGGTGTGGGCCAGGTAGACTTCAACACCTTGAAGTCCATTG  
GGAGTGAGCCCCATGAGGACCATGTCTTCTTGTGGCCAATTTAGCCAGATTGAGACGCTGACCTCCGTGTTCC  
AGAAGAAGTTGTGCACGGCCACATGTGCAGCACCCCTGGAGCATAACTGTGCCCACTTCTGCATCAACATCCCTG  
GCTCATACGTCTGCAGGTGCAAACAAGGCTACATTCTCAACTCGGATCAGACGACTTGCAGAATCCAGGATCTGT  
GTGCCATGGAGGACCACAACCTGTGAGCAGCTCTGTGTGAATGTGCCGGGCTCCTTCGTCTGCCAGTGCTACAGTG  
GCTACGCCCTGGCTGAGGATGGGAAGAGGTGTGTGCTGTGGACTACTGTGCCCTCAGAAAACCACGGATGTGAAC  
ATGAGTGTGTAAATGCTGATGGCTCCTACCTTTGCCAGTGCCATGAAGGATTGTCTTTAACCAGATGAAAAAA  
CGTGCAAGGATCAACTACTGTGCACTGAACAAACCGGGCTGTGAGCATGAGTGCCTCAACATGGAGGAGAGCT  
ACTACTGCCGCTGCCACCGTGGCTACACTCTGGACCCCAATGGCAAACCTGCAGCCGAGTGGACCACTGTGCAC  
AGCAGGACCATGGCTGTGAGCAGCTGTGTCTGAACACGGAGGATTCTTCGTCTGCCAGTGCTCAGAAGGCTTCC  
TCATCAACGAGGACCTCAAGACCTGCTCCCGGTGGATTACTGCCTGCTGAGTGACCATGGTTGTGAATACTCCT  
GTGTCAACATGGACAGATCCTTTGCCCTGTAGTGTCTGAGGGACAGTGCTCCGACGATGGGAAGACGTGTG  
CAAAATTGGACTCTTGTGCTCTGGGGGACCACGGTTGTGAACATTCGTGTGTAAAGCAGTGAAGATTCTGTTGTG  
GCCAGTGCTTTGAAGGTTATATACTCCGTGAAGATGGAAAAACCTGCAGAAGGAAAGATGTCTGCCAAGCTATAG  
ACCATGGCTGTGAACACATTTGTGTGAACAGTGACGACTCATACAGTGCGAGTGCTTGGAGGGATTCCGGCTCG  
CTGAGGATGGGAAACGCTGCCGAAGGAAGGATGTCTGCAAATCAACCCACCATGGCTGCCAACACATTTGTGTTA  
ATAATGGGAATTCCTACATCTGCAAATGCTCAGAGGGATTTGTTCTAGCTGAGGACGGAAGACGGTGCAAGAAAT  
GCACCTGAAGGCCCAATTGACCTGGTCTTTGTGATCGATGGATCCAAGAGTCTTGGAGAAGAGAATTTTGAGGTG  
TGAAGCAGTTTGTCACTGGAATTATAGATTCTTTGACAATTTCCCCCAAAGCCGCTCGAGTGGGGCTGCTCCAGT  
ATTCCACACAGGTCCACACAGAGTTCACTCTGAGAAACTTCAACTCAGCCAAAGACATGAAAAAAGCCGTGGCCC  
ACATGAAATACATGGGAAGGGCTCTATGACTGGGCTGGCCCTGAAACACATGTTTGAGAGAAGTTTACCCAAG  
GAGAAGGGGCCAGGCCCCCTTTCCACAAGGGTGCCGAGAGCAGCCATTGTGTTTACCGACGGACGGGCTCAGGATG  
ACGTCTCCGAGTGGGCCAGTAAAGCCAAGGCCAATGGTATCACTATGTATGCTGTTGGGGTAGGAAAAGCCATTG  
AGGAGGAACATAAAGAGATTGCCCTCTGAGCCCAACAACAAGCATCTCTTCTATGCCGAAGACTTCAACACAATGG  
ATGAGATAAGTGAAAAACTCAAGAAAGGCATCTGTGAAGCTCTAGAAGACTCCGATGGAAGACAGGACTCTCCAG  
CAGGGGAACCTGCCAAAAACGGTCCAACAGCCAACAGAATCTGAGCCAGTCACCATAAATATCCAAGACCTACTTT  
CCTGTTCTAATTTTGCAGTGCAACACAGATATCTGTTTGAAGAAGACAATCTTTTACGGTCTACACAAAAGCTTT  
CCCATTCAACAAAACCTTCAGGAAGCCCTTTGGAAGAAAAACACGATCAATGCAAATGTGAAAACTTATAATGT  
TCCAGAACCTTGCAAACGAAGAAGTAAGAAAATTAACACAGCGCTTAGAAGAAATGACACAGAGAATGGAAGCCC  
TGAAAAATCGCCTGAGATACAGATGAAGATTAGAAATCGCGACACATTTGTAGTCATTGTATCACGGATTACAAT  
GAACGCAGTGACAGAGCCCCAAAGCTCAGGCTATTGTTAAATCAATAATGTTGTGAAGTAAAAACAATCAGTACTGA  
GAAACCTGGTTTGGCACAGAACAAAGACAAGAAGTATACACTAAGTGTATAAATTTATCTAGGAAAAAAATCCT  
TCAGAATTCCTAAGATGAATTTACCAGGTGAGAATGAATAAGCTATGCAAGGTATTTTGTAAATATACTGTGGACAC  
AATTGCTTCTGCCTCATCTGCCTTAGTGTGAATCTCATTTGACTATACGATAAAGTTTGACAGTCTTACTT  
CTGTAGAACACTGGCCATAGGAAATGCTGTTTTTTTGTACTGGACTTTACCTTGATATATGTATATGGATGTATG  
CATAAAATCATAGGACATATGTACTTGTGGAACAAGTTGGATTTTTTATACAATATTAATAATTCACCACTTCAG

Sequence

## FIGURE 15

MEKMLAGCFLILGQIVLLPAEARERSRGRSISRGRHARTHPTALLESSCENKRADLVFII  
DSSRSVNTHDYAKVKEFIVDILQFLDIGPDVTRVGLLQYGSTVKNEFSLKTFKRKSEVERAV  
KMRHLSTGTMTGLAIQYALNIAFSEAEGARPLRENVPRVIMIVTDGRPQDSVAEVAAKARD  
TGILIFAIGVGQVDFNTLKSIGSEPHEDHVFLVANFSQIETLTSVFQKKLCTAHMCSTLEHN  
CAHFCINIPGSYVCRCKQGYILNSDQTTCRIQDLCAMEDHNCEQLCVNVPGSFVCQCYSGYA  
LAEDGKRCVAVDYCASENHGCEHECVNADGSYLCQCHEGFALNPDEKTCTRINYCALNKP GC  
EHECVNMEESYYCRCHRGYTLDPNGKTC SRVDHCAQQDHGCEQLCLNTEDSFVCQCSEGFLI  
NEDLKTC SRVDYCLLSDHGCEYSCVNMDRSFACQCPEGHVLRSDGKTC AKLDSCALGDHGCE  
HSCVSSDSFVCQCFEGYILREDGKTCRRKDVCQAIDHGCEHICVNSDDSYTCECLEGFRLA  
EDGKRCRRKDVCCKSTHHGCEHICVNNGNSYICKCSEGFVLAEDGRRCKKCTEGPIDLVFVID  
GSKSLGEENFEVVKQFVTGIIDSLTISPKAARVGLLOYSTQVHTEFTLRNFNSAKDMKKAVA  
HMKYMGKGSMTGLALKHMFERSFTQEGGARPLSTRVPRAAIVFTDGRAQDDVSEWASKAKAN  
GITMYAVGVGKAIEEELQEIASEPTNKHLYFAEDFSTMDIEISEKLKKGICEALEDSDGRQDS  
PAGELPKTVQQPTESEPVTINIQDLLSCSNFAVQHRYLFEEDNLLRSTQKLSHSTKPSGSPL  
EEKHDQCKCENLIMFQNLANEEVRKLTQRLEEMTQORMEALENRLRYR

**Signal peptide:**

amino acids 1-23

**N-glycosylation site.**

amino acids 221-225

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 115-119, 606-610, 892-896

**Casein kinase II phosphorylation site.**

amino acids 49-53, 118-122, 149-153, 176-180, 223-227, 243-247,  
401-405, 442-446, 501-505, 624-628, 673-677, 706-710, 780-784,  
781-785, 819-823, 866-870

**N-myristoylation site.**

amino acids 133-139, 258-264, 299-305, 340-346, 453-459, 494-500,  
639-645, 690-696, 752-758, 792-798

**Amidation site.**

amino acids 314-318, 560-564, 601-605

**Aspartic acid and asparagine hydroxylation site.**

amino acids 253-265, 294-306, 335-347, 376-388, 417-423, 458-464,  
540-546, 581-587

## FIGURE 16

GGAGCCGCCCTGGGTGTCAGCGGCTCGGCTCCCGCGCACGCTCCGGCCGTGCGCGAGCCTCG  
GCACCTGCAGGTCCGTGCGTCCCGCGGCTGGCGCCCCCTGACTCCGTCCCGGCCAGGGAGGGC  
**CATG**ATTTCCTCCCGGGGCCCCCTGGTGACCAACTTGCTGCGGTTTTTGTTCCTGGGGCTGA  
GTGCCCTCGCGCCCCCTCGCGGGCCCAGCTGCAACTGCACCTGCCCGCCAACCGGTTGCAG  
GCGGTGGAGGGAGGGGAAGTGGTGCTTCCAGCGTGGTACACCTTGACGCGGGAGGTGTCTTC  
ATCCCAGCCATGGGAGGTGCCCTTTGTGATGTGGTTCTTCAAACAGAAAGAAAAGGAGGATC  
AGGTGTTGTCCTACATCAATGGGGTCACAACAAGCAAACCTGGAGTATCCTTGGTCTACTCC  
ATGCCCTCCCGGAACCTGTCCCTGCGGCTGGAGGGTCTCCAGGAGAAAGACTCTGGCCCCTA  
CAGCTGCTCCGTGAATGTGCAAGACAAACAAGGCAAATCTAGGGGCCACAGCATCAAAACCT  
TAGAACTCAATGTACTGGTTCCTCCAGCTCCTCCATCCTGCCGTCTCCAGGGTGTGCCCCAT  
GTGGGGGCAAACGTGACCCTGAGCTGCCAGTCTCCAAGGAGTAAGCCCGCTGTCCAATACCA  
GTGGGATCGGCAGCTTCCATCCTTCCAGACTTTCTTTGCACCAGCATTAGATGTCATCCGTG  
GGTCTTTAAGCCTCACCAACCTTTCGTCTTCCATGGCTGGAGTCTATGTCTGCAAGGCCAC  
AATGAGGTGGGCACTGCCCAATGTAATGTGACGCTGGAAGTGAGCACAGGGCCTGGAGCTGC  
AGTGGTTGCTGGAGCTGTTGTGGGTACCCTGGTTGGACTGGGGTTGCTGGCTGGGCTGGTCC  
TCTTGTACCACCGCCGGGGCAAGGCCCTGGAGGAGCCAGCCAATGATATCAAGGAGGATGCC  
ATTGCTCCCCGGACCCTGCCCTGGCCCAAGAGCTCAGACACAATCTCCAAGAATGGGACCCT  
TTCCTCTGTACCTCCGCACGAGCCCTCCGGCCACCCCATGGCCCTCCCAGGCCTGGTGCAT  
TGACCCCCACGCCAGTCTCTCCAGCCAGGCCCCGCCCCACCAAGACTGCCACGACAGAT  
GGGGCCCACCCTCAACCAATATCCCCCATCCCTGGTGGGGTTTTCTTCTCTGGCTTGAGCCG  
CATGGGTGCTGTGCCTGTGATGGTGCCTGCCAGAGTCAAGCTGGCTCTCTGGTAT**TG**ATGAC  
CCCACCACTCATTGGCTAAAGGATTTGGGGTCTCTCCTTCCCTATAAGGGTCACCTCTAGCAC  
AGAGGCCTGAGTCATGGGAAAGAGTCACACTCCTGACCCTTAGTACTCTGCCCCACCTCTC  
TTTACTGTGGGAAAACCATCTCAGTAAGACCTAAGTGTCCAGGAGACAGAAGGAGAAGAGGA  
AGTGGATCTGGAATTGGGAGGAGCCTCCACCCACCCCTGACTCCTCCTTATGAAGCCAGCTG  
CTGAAATTAGCTACTACCAAGAGTGAGGGGCAGAGACTTCCAGTCACTGAGTCTCCAGGC  
CCCCTTGATCTGTACCCACCCCTATCTAACACCACCCTTGGCTCCCACTCCAGCTCCCTGT  
ATTGATATAACCTGTGAGGCTGGCTTGGTTAGGTTTTACTGGGGCAGAGGATAGGGAATCTC  
TTATTAAACTAACATGAAATATGTGTTGTTTTCATTTGCAAATTTAAATAAAGATACATAA  
TGTTTGTATGAAAAA

## **FIGURE 17**

MISLPGPLVTNLLRFLFLGLSALAPPSRAQLQLHLPANRLQAVEGGEVVLPAWYTLHGEVSS  
SQPWEVPFVMWFFKQKEKEDQVLSYINGVTTSKPGVSLVYSMPSRNLSLRLEGLQEKDSGPY  
SCSVNVQDKQGKSRGHSIKTLELNVLVPPAPPSCRLQGVPHVGANVTLSQCSPRSKPAVQYQ  
WDRQLPSFQTFAPALDVIRGSLSLTNLSSSMAGVYVCKAHNEVGTAQCNVTLEVSTGPGAA  
VVAGAVVGTLVGLGLLAGLVLLYHRRGKALEEPANDIKEDAIAPRTLWPWKSSDTISKNGTL  
SSVTSARALRPPHGP RP GALTP T PSLSSQALPSPRLPTTDGAHPQPISPIPGGVSSSGLSR  
MGAVPVMVPAQSQAGSLV

### **Signal peptide:**

amino acids 1-29

### **Transmembrane domain:**

amino acids 245-267

### **N-glycosylation site.**

amino acids 108-112, 169-173, 213-217, 236-240, 307-311

### **N-myristoylation site.**

amino acids 90-96, 167-173, 220-226, 231-237, 252-258, 256-262,  
262-268, 308-314, 363-369, 364-370

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 164-175



## **FIGURE 19**

MKRLPLLVVFSTLLNCSYTONCTKTPCLPNAKCEIRNGIEACYCNMGFSGNGVTICEDDNEC  
GNLTQSCGENANCTNTEGSYYCMCVPGFRSSSNQDRFITNDGTVCIENVNANCHLDNVCIAA  
NINKTLTKIRSIKEPVALLQEVYRNSVTDLSPTDIITYIEILAESSSLLGYKNNTISAKDTL  
SNSTLTFVKTNNFVQORDTFVVWDKLSVNHRRTHLTKLMHTVEQATLRISQSFQKTTEFDT  
NSTDIALKVFFFDSYNMKHIHPHMNDGDYINIFPKRKAAYDSNGNVAVAFLLYYKSIGPLLS  
SSDNFLLKPQNYDNSEEEERVISSVISVSMSSNPPTLYELEKITFTLSHRKVTDRYRSLCAF  
WNYSPTMNGSWSSEGCETYSNETHTSCRCNHLTHFAILMSSGPSIGIKDYNILTRITQLG  
IIISLICLAICIFTFWFFSEIQSTRTTIHKNLCCSLFLAELVFLVGINTNTNKLFCsiiAGL  
LHYFFLAFAWMCIIEGHLVIVGVYINKGFLHKNFYIFGYLSPAVVVGFSALGYRYYGT  
TKVCWLSTENNFIFWFIGPACLIILVNLLAFGVIIYKVFRHTAGLKPEVSCFENIRSCARGA  
LALLFLLGTTWIFGVLVHVVHASVVTAYLFTVSNAFQGMFIFLFLCVLSRKIQEYYRLFKNV  
PCCFGCLR

### **Signal peptide:**

amino acids 1-19

### **Transmembrane domain:**

amino acids 430-450, 465-486, 499-513, 535-549, 573-593, 619-636,  
648-664

### **N-glycosylation site.**

amino acids 15-19, 21-25, 64-68, 74-78, 127-131, 177-181,  
188-192, 249-253, 381-385, 395-399

### **Glycosaminoglycan attachment site.**

amino acids 49-53

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 360-364

### **Casein kinase II phosphorylation site.**

amino acids 54-58, 68-72, 76-80, 94-98, 135-139, 150-154,  
155-159, 161-165, 181-185, 190-194, 244-248, 310-314, 325-329,  
346-350, 608-612

### **Tyrosine kinase phosphorylation site.**

amino acids 36-44, 669-677, 670-678

### **N-myristoylation site.**

amino acids 38-44, 50-56, 52-58, 80-86, 382-388, 388-394,  
434-440, 480-486, 521-527

### **Aspartic acid and asparagine hydroxylation site.**

amino acids 75-87



## FIGURE 20

TGGAAACATATCCTCCCTCATATGAATATGGATGGAGACTACATAAATATATTTCAAAGNG  
AAAAGCCGGCATATGGATTCAAATGGCAATGTTGCAGTTGCATTTTTATATTATAAGAGTAT  
TGGTCCCTTTGCTTTTCATCATCTGACAACCTTCTTATTGAAACCTCAAATTTATGATAATTCT  
GAAGAGGAGGAAAGAGTCATATCTTCAGTAATTTTCAGTCTCAATGAGCTCAAACCCACCCAC  
ATTATATGAACTTGAAAAATAACATTTACATTAAGTCATCGAAAGGTCACAGATAGGTATA  
GGAGTCTATGTGGCATTGTTGGAATACTCACCTGATACCATGAATGGCAGCTGGTCTTCAGAG  
GGCTGTGAGCTGACATACTCAAATGAGACCCACACCTCATGCCGCTGTAATCACCTGACACA  
TTTTGCAATTTTGATGTCCTCTGGTCCTTCCATTGGTATTAAAGATTATAATATTCTTACAA  
GGATCACTCAACTAGGAATAATTATTTCACTGATTTGTCTTGCCATATGCATTTTTACCTTC  
TGGTTCTTCAGTGAAATTCAAAGCACCAGGA

ref. 20 00000000



## FIGURE 21

GCTCCCAGCCAAGAACCTCGGGGCCGCTGCGCGGTGGGGAGGAGTTCCCCGAAACCCGGGCCG  
CTAAGCGAGGCCTCCTCCTCCCGCAGATCCGAACGGCCTGGGCGGGGTACCCCGGCTGGGA  
CAAGAAGCCGCGCCTGCCTGCCCGGGCCCGGGGAGGGGGCTGGGGCTGGGGCCGGAGGCGG  
GGTGTGAGTGGGTGTGTGCGGGGGGCGGAGGCTTGATGCAATCCCGATAAGAAATGCTCGGG  
TGTCTTGGGCACCTACCCGTGGGGCCCCGTAAGGCGCTACTATATAAGGCTGCCGGCCCCGAG  
CCGCGCGCGCGTCAGAGCAGGAGCGCTGCGTCCAGGATCTAGGGCCACGACCATCCCAACCC  
GGCACTCACAGCCCCGAGCGCATCCCGGTGCGCGCCAGCCTCCCGCACCCCCATCGCCGG  
AGCTGCGCCGAGAGCCCCAGGGAGGTGCCATGCGGAGCGGGTGTGTGGTGGTCCACGTATGG  
ATCCTGGCCGGCCTCTGGCTGGCCGTGGCCGGGCGCCCCCTCGCCTTCTCGGACGCGGGGCC  
CCACGTGCACTACGGCTGGGGCGACCCCATCCGCCTGCGGCACCTGTACACCTCCGGCCCCC  
ACGGGCTCTCCAGCTGCTTCTGCGCATCCGTGCCGACGGCGTCGTGGACTGCGCGCGGGGC  
CAGAGCGCGCACAGTTTGCTGGAGATCAAGGCAGTCGCTCTGCGGACCGTGGCCATCAAGGG  
CGTGACAGCGTGCGGTACCTCTGCATGGGGCGCCGACGGCAAGATGCAGGGGCTGCTTCAGT  
ACTCGGAGGAAGACTGTGCTTTCGAGGAGGAGATCCGCCCAGATGGCTACAATGTGTACCGA  
TCCGAGAAGCACCGCCTCCCGGTCTCCCTGAGCAGTGCCAAACAGCGGCAGCTGTACAAGAA  
CAGAGGCTTTCTTCCACTCTCTCATTTTCTGCCCATGCTGCCCATGGTCCCAGAGGAGCCTG  
AGGACCTCAGGGGCCACTTGGAATCTGACATGTTCTCTTCGCCCTGGAGACCGACAGCATG  
GACCCATTTGGGCTTGTCACCGGACTGGAGGCCGTGAGGAGTCCCAGCTTTGAGAAGTAACT  
GAGACCATGCCCCGGGCCTCTTCACTGCTGCCAGGGGCTGTGGTACCTGCAGCGTGGGGGACG  
TGCTTCTACAAGAACAGTCCTGAGTCCACGTTCTGTTTtagCTTTAGGAAGAAACATCTAGAA  
GTTGTACATATTCAGAGTTTTCCATTGGCAGTGCCAGTTTCTAGCCAATAGACTTGTCTGAT  
CATAACATTGTAAGCCTGTAGCTTGCCCAGCTGCTGCCTGGGCCCCCATTCTGCTCCCTCGA  
GGTTGCTGGACAAGCTGCTGCACTGTCTCAGTTCTGCTTGAATACCTCCATCGATGGGGAAC  
TCACTTCCTTTGGAAAAATTCTTATGTCAAGCTGAAATTCTCTAATTTTTTCTCATCACTTC  
CCCAGGAGCAGCCAGAAGACAGGCAGTAGTTTTAATTTAGGAACAGGTGATCCACTCTGTA  
AAACAGCAGGTAAATTTCACTCAACCCCATGTGGGAATTGATCTATATCTCTACTTCCAGGG  
ACCATTTGCCCTTCCCAAATCCCTCCAGGCCAGAACTGACTGGAGCAGGCATGGCCCCACCAG  
GCTTCAGGAGTAGGGGAAGCCTGGAGCCCCACTCCAGCCCTGGGACAACTTGAGAATTCCCC  
CTGAGGCCAGTTCTGTCTATGGATGCTGTCTGAGAATAACTTGCTGTCCCGGTGTCACCTGC  
TTCCATCTCCCAGCCCACCAGCCCTCTGCCCACCTCACATGCCTCCCCATGGATTGGGGCCT  
CCCAGGCCCCCCCACCTTATGTCAACCTGCACTTCTTGTTCAAAAATCAGGAAAAGAAAAGAT  
TTGAAGACCCCAAGTCTTGTCAATAACTTGCTGTGTGGAAGCAGCGGGGGAAGACCTAGAAC  
CCTTTCCCCAGCACTTGGTTTTTCCAACATGATATTTATGAGTAATTTATTTTGATATGTACA  
TCTCTTATTTTCTTACATTATTTATGCCCCAAATTATATTTATGTATGTAAGTGAGGTTTG  
TTTTGTATATTAAATGGAGTTTGT

## **FIGURE 22**

MRSGCVVHVWILAGLWLAVAGRPLAFSDAGPHVHYGWDPIRLRHLYTSGPHGLSSCFLRI  
RADGVVDCARGQSAHSLLEIKAVALRTVAIKGVH SVRYLCMGADGKMQGLLQYSEEDCAFEE  
EIRPDGYNVYRSEKHRLPVSLSSAKQRQLYKNRGFLPLSHFLPMLPMVPEEPEDLRGHLESD  
MFSSPLETDSMDPFGLVTGLEAVRSPSF EK

### **Signal peptide:**

amino acids 1-22

### **Casein kinase II phosphorylation site.**

amino acids 78-82, 116-120, 190-194, 204-208

### **N-myristoylation site.**

amino acids 15-21, 54-60, 66-72, 201-207

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 48-59

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## **FIGURE 23**

CCCAGAAAGTTCAAGGGCCCCCGGCCTCCTGCGCTCCTGCCGCCGGGACCCTCGACCTCCTCA  
GAGCAGCCGGCTGCCGCCCCGGGAAGATGGCGAGGAGGAGCCGCCACCGCCTCCTCCTGCTG  
CTGCTGCGCTACCTGGTGGTTCGCCCTGGGCTATCATAAGGCCTATGGGTTTTCTGCCCCAAA  
AGACCAACAAGTAGTCACAGCAGTAGAGTACCAAGAGGCTATTTTAGCCTGCAAAACCCCAA  
AGAAGACTGTTTTCTCCAGATTAGAGTGGAAGAACTGGGTTCGGAGTGTCTCCTTTGTCTAC  
TATCAACAGACTCTTCAAGGTGATTTTAAAAATCGAGCTGAGATGATAGATTTCAATATCCG  
GATCAAAAATGTGACAAGAAGTGATGCGGGGAAATATCGTTGTGAAGTTAGTGCCCCATCTG  
AGCAAGGCCAAAACCTGGAAGAGGATACAGTCACTCTGGAAGTATTAGTGGCTCCAGCAGTT  
CCATCATGTGAAGTACCCTCTTCTGCTCTGAGTGGAAGTGTGGTAGAGCTACGATGTCAAGA  
CAAAGAAGGGAATCCAGCTCCTGAATACACATGGTTTAAGGATGGCATCCGTTTGCTAGAAA  
ATCCCAGACTTGGCTCCCAAAGCACCAACAGCTCATAACAATGAATACAAAACTGGAAGT  
CTGCAATTTAATACTGTTTCCAACTGGGACACTGGAGAATATTCTGTGAAGCCCGCAATTC  
TGTTGGATATCGCAGGTGTCCTGGGAAACGAATGCAAGTAGATGATCTCAACATAAGTGGCA  
TCATAGCAGCCGTAGTAGTTGTGGCCTTAGTGATTTCCGTTTGTGGCCTTGGTGTATGCTAT  
GCTCAGAGGAAAGGCTACTTTTCAAAGAAACCTCCTTCCAGAAGAGTAATTCTTCATCTAA  
AGCCACGACAATGAGTGAAAATGTGCAGTGGCTCACGCCTGTAATCCCAGCACTTTGGAAGG  
CCGCGGCGGGCGGATCACGAGGTCAGGAGTTCTAGACCAGTCTGGCCAATATGGTGAAACCC  
CATCTCTACTAAAATACAAAAATTAGCTGGGCATGGTGGCATGTGCCTGCAGTTCCAGCTGC  
TTGGGAGACAGGAGAATCACTTGAACCCGGGAGGCGGAGGTTGCAGTGAGCTGAGATCACGC  
CACTGCAGTCCAGCCTGGGTAACAGAGCAAGATTCATCTCAAAAATAAAATAAATAAATA  
AATAAATACTGGTTTTTACCTGTAGAATTCTTACAATAAATATAGCTTGATATTC

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## **FIGURE 24**

MARRSRHRLLLLLLRYLVVALGYHKAYGFSAPKDQQVVTAVEYQEAILACKTPKKTVSSRLE  
WKKLGRSVSFVYYQQTLQGDFKNRAEMIDFNIRIKNVTRSDAGKYRCEVSAPSEQGQONLEED  
TVTLEVLVAPAVPSCCEVPSSALSGTVVELRCQDKEGNPAPEYTWFKDGIRLLENPRLGSQST  
NSSYTMNTKTGTLQFNTVSKLDTGEYSCEARNVSGYRRCPGKRMQVDDLNIISGIIAAVVVVA  
LVISVCGLGVCYAQRKGYFSKETSFOKSNSSSKATTMSENVQWLTPVIPALWCAAAGGSRGQEF

### **Signal peptide:**

amino acids 1-20

### **Transmembrane domain:**

amino acids 130-144, 238-258

### **N-glycosylation site.**

amino acids 98-102, 187-191, 236-240, 277-281

### **Casein kinase II phosphorylation site.**

amino acids 39-43, 59-63, 100-104, 149-153, 205-209, 284-288

### **N-myristoylation site.**

amino acids 182-188, 239-245, 255-261, 257-263, 305-311

### **Amidation site.**

amino acids 226-230

FIGURE 24

## FIGURE 25

GACATCGGAGGTGGGCTAGCACTGAACTGCTTTTCAAGACGAGGAAGAGGAGGAGAAAGAG  
AAAGAAGAGGAAGATGTTGGGCAACATTTATTTAACATGCTCCACAGCCCGGACCCTGGCAT  
CATGCTGCTATTCTCTGCAAATACTGAAGAAGCATGGGATTTAAATATTTTACTTCTAAATAA  
ATGAATTACTCAATCTCCTATGACCATCTATACATACTCCACCTTCAAAAAGTACATCAATA  
TTATATCATTAAAGGAAATAGTAACCTTCTCTTCTCCAATATGCATGACATTTTGGACAATG  
CAATTGTGGCACTGGCACTTATTTTCAGTGAAGAAAACTTTGTGGTTCTATGGCATTTCATCA  
TTTGACAAATGCAAGCATCTTCCTTATCAATCAGCTCCTATTGAACCTTACTAGCACTGACTG  
TGGAATCCTTAAGGGCCCATTACATTTCTGAAGAAGAAAGCTAAGATGAAGGACATGCCACT  
CCGAATTCATGTGCTACTTGGCCTAGCTATCACTACACTAGTACAAGCTGTAGATAAAAAAG  
TGGATTGTCCACGGTTATGTACGTGTGAAATCAGGCCTTGGTTTACACCCAGATCCATTTAT  
ATGGAAGCATCTACAGTGGATTGTAATGATTTAGGTCTTTTAACTTTCCCAGCCAGATTGCC  
AGCTAACACACAGATTCTTCTCCTACAGACTAACAATATTGCAAAAATTGAATACTCCACAG  
ACTTTCCAGTAAACCTTACTGGCCTGGATTTATCTCAAAACAATTTATCTTCAGTCACCAAT  
ATTAATGTAAAAAAGATGCCTCAGCTCCTTTCTGTGTACCTAGAGGAAAACAACTTACTGA  
ACTGCCTGAAAAATGTCTGTCCGAAGTGAAGCACTTACAAGAACTCTATATTAATCACAAC  
TGCTTTCTACAATTTACCTGGAGCCTTTATTGGCCTACATAATCTTCTTCGACTTCATCTC  
AATTCAAATAGATTGCAGATGATCAACAGTAAGTGGTTTGATGCTCTTCCAAATCTAGAGAT  
TCTGATGATTGGGGAAAATCCAATTATCAGAATCAAAGACATGAACCTTTAAGCCTCTTATCA  
ATCTTCGCAGCCTGGTTATAGCTGGTATAAACCTCACAGAAATACCAGATAACGCCTTGGTT  
GGACTGGAAAACCTTAGAAAGCATCTCTTTTACGATAACAGGCTTATTAAAGTACCCCATGT  
TGCTCTTCAAAAAGTTGTAAATCTCAAATTTTGGATCTAAATAAAAAATCCTATTAATAGAA  
TACGAAGGGGTGATTTTAGCAATATGCTACACTTAAAAGAGTTGGGGATAAATAATATGCCT  
GAGCTGATTTCCATCGATAGTCTTGCTGTGGATAACCTGCCAGATTTAAGAAAAATAGAAGC  
TACTAACAAACCCTAGATTGTCTTACATTCACCCCAATGCATTTTTCAGACTCCCCAAGCTGG  
AATCACTCATGCTGAACAGCAATGCTCTCAGTGCCCTGTACCATGGTACCATTGAGTCTCTG  
CCAAACCTCAAGGAAATCAGCATAACAGTAACCCCATCAGGTGTGACTGTGTGCATCCGTTG  
GATGAACATGAACAAAACCAACATTCGATTTCATGGAGCCAGATTCAGTGTGCGTGGACC  
CACCTGAATTCAGGTGAGATGTTTCGGCAAGTGCATTTTCAGGGACATGATGGAAATTTGT  
CTCCCTCTTATAGCTCCTGAGAGCTTTCTCTTCTAATCTAAATGTAGAAGCTGGGAGCTATGT  
TTCTTTTCACTGTAGAGCTACTGCAGAACACAGCCTGAAATCTACTGGATAACACCTTCTG  
GTCAAAAACCTTTGCCTAATACCCTGACAGACAAGTTCTATGTCCATTCTGAGGGAACACTA  
GATATAAATGGCGTAACTCCCAAAGAAGGGGTTTATATACTTGTATAGCAACTAACCTAGT  
TGGCGCTGACTTGAAGTCTGTTATGATCAAAGTGGATGGATCTTTTCCACAAGATAACAATG  
GCTCTTTGAATATTAATAAAGAGATATTGAGGCAATTCAGTTTTGGTGTCTTGGAAAGCA  
AGTTCTAAAATTCTCAAATCTAGTGTAAATGGACAGCCTTTGTCAAGACTGAAAATTCTCA  
TGCTGCGCAAAGTGCTCGAATACCATCTGATGTCAAGGTATATAATCTTACTCATCTGAATC  
CATCAACTGAGTATAAAAATTTGTATTGATATTCCCACCATCTATCAGAAAAACAGAAAAAA  
TGTGTAAATGTCACCACCAAGGTTTGCACCCTGATCAAAAAGAGTATGAAAAGAATAATAC  
CACAACACTTATGGCCTGTCTTGGAGGCCTTCTGGGGATTATTGGTGTGATATGTCTTATCA  
GCTGCCTCTCTCCAGAAATGAACTGTGATGGTGGACACAGCTATGTGAGGAATTACTTACAG  
AAACCAACCTTTGCATTAGGTGAGCTTTATCCTCCTCTGATAAATCTCTGGGAAGCAGGAAA  
AGAAAAAAGTACATCACTGAAAGTAAAAGCAACTGTTATAGGTTTACCAACAAATATGTCTT  
AAAAACCACCAAGGAAACCTACTCCAAAAATGAAC

## **FIGURE 26**

MKDMPLRIHVLLGLAITTLVQAVDKKVDPCRLCTCEIRPWFTPRSIYMEASTVDCNDLGLLT  
FPARLPANTQILLQTNNAKIEYSTDFVNLTGLDLSQNNLSSVTNINVKKMPQLLSVYLE  
ENKLTPEKCLSELNLQELYINHNLLSTISPGAFIHLNLLRLHLNSNRLQMINSKWFDA  
LPNLEILMIGENPIIRIKDMNFKPLINLRSLVIAGINLTEIPDNALVGLENLESISFYDNRL  
IKVPHVALQKVNLKFLDLNKNPINRIRRGDFSNNMLHLKELGINNMPELISIDSLAVDNLDP  
LRKIEATNNPRLSYIHPNAFFRLPKLESLMLNSNALSALYHGTIESLPNLKEISIHSPNIRC  
DCVIRWMNMNKTNIRFMEPDSLFCVDPPEFQGNVRQVHFRDMMEICLPLIAPESFPSNLNV  
EAGSYVSFHCRTAEAPQPEIYWITPSGQKLLPNTLTDFYVHSEGTLDINGVTPKEGGLYTC  
IATNLVGADLKSVMIKVDGSFPQDNNGSLNIKIRDIQANSVLVSWKASSKILKSSVKWTAFAV  
KTENSHAAQSARIPSDVKVYNLTHLNPSTEYKICIDIPTIYQKNRKKCVNVTTKGLHPDQKE  
YEKNNTTTLMACLGGLLGIIGVICLISCLSPENMCDGGHSYVRNYLQKPTFALGELYPPLIN  
LWEAGKEKSTSLKVKATVIGLPTNMS

### **Signal sequence:**

amino acids 1-22

### **Transmembrane domain:**

amino acids 633-650

### **N-glycosylation site.**

amino acids 93-97, 103-107, 223-227, 382-386, 522-526, 579-583,  
608-612, 624-628, 625-629

### **Casein kinase II phosphorylation site.**

amino acids 51-55, 95-99, 242-246, 468-472, 487-491

### **Tyrosine kinase phosphorylation site.**

amino acids 570-579

### **N-myristoylation site.**

amino acids 13-19, 96-102, 158-164, 221-227, 352-358, 437-443,  
491-497, 492-498, 634-640, 702-708

### **Cell attachment sequence.**

amino acids 277-280

FIGURE 26

## **FIGURE 27**

GCCCCGGGACTGGCGCAAGGTGCCCAAGCAAGGAAAGAAATAATGAAGAGACACATGTGTAG  
CTGCAGCCTTTTGAAACACGCAAGAAGGAAATCAATAGTGTGGACAGGGCTGGAACCTTTAC  
CACGCTTGTTGGAGTAGATGAGGAATGGGCTCGTGATTATGCTGACATTCCAGCATGAATCT  
GGTAGACCTGTGGTTAACCCGTTCCCTCTCCATGTGTCTCCTCCTACAAAGTTTTGTTCTTA  
TGATACTGTGCTTTCATTCTGCCAGTATGTGTCCCAAGGGCTGTCTTTGTTCTTCCTCTGGG  
GGTTTAAATGTCACCTGTAGCAATGCAAATCTCAAGGAAATACCTAGAGATCTTCCTCCTGA  
AACAGTCTTACTGTATCTGGACTCCAATCAGATCACATCTATTCCCAATGAAATTTTTTAAGG  
ACCTCCATCAACTGAGAGTTCTCAACCTGTCCAAAAATGGCATTGAGTTTATCGATGAGCAT  
GCCTTCAAAGGAGTAGCTGAAACCTTGCAGACTCTGGACTTGTCCGACAATCGGATTCAAAG  
TGTGCACAAAAATGCCTTCAATAACCTGAAGGCCAGGGCCAGAATTGCCAACAACCCCTGGC  
ACTGCGACTGTACTCTACAGCAAGTTCTGAGGAGCATGGCGTCCAATCATGAGACAGCCCAC  
AACGTGATCTGTAAAACGTCCGTGTTGGATGAACATGCTGGCAGACCATTCTCAATGCTGC  
CAACGACGCTGACCTTTGTAACTCCCTAAAAAACTACCGATTATGCCATGCTGGTCACCA  
TGTTTGGCTGGTTCACCTATGGTGATCTCATATGTGGTATATTATGTGAGGCAAAATCAGGAG  
GATGCCCGGAGACACCTCGAATACTTGAAATCCCTGCCAAGCAGGCAGAAGAAAGCAGATGA  
ACCTGATGATATTAGCACTGTGGTATAGTGTCCAACTGACTGTCATTGAGAAAGAAAGAAA  
GTAGTTTGCGATTGCAGTAGAAATAAGTGGTTTACTTCTCCCATCCATTGTAAACATTTGAA  
ACTTTGTATTTTCACTTTTTTTTTGAATTATGCCACTGCTGAACTTTTAACAAACACTACAACA  
TAAATAATTTGAGTTTAGGTGATCCACCCCTTAATTGTACCCCGATGGTATATTTCTGAGT  
AAGCTACTATCTGAACATTAGTTAGATCCATCTCACTATTTAATAATGAAATTTATTTTTTTT  
AATTTAAAAGCAAATAAAAGCTTAACTTTGAACCATGGGAAAAAAAAAAAAAAAAAAAAAAAAACA

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## **FIGURE 28**

MNLVDLWLTRSLMCLLLQSFLMILCFHSASMCPKGCLCSSSGGLNVTC SNANLKEI PRDL  
PPETVLLYLDSNQITSIPNEIFKDLHQLRVLNLSKNGIEFIDEHAFKGVAETLQTLDSLSDNR  
IQSVHKNAFNNL KARARIANNPW HCDCTLQQVLRSMASNHETAHNVICKTSVLDEHAGRPFL  
NAANDADLCNLPKKT TDYAMLVTMFGWFTMVISYVVYYVRQNQEDARRHLEYLKSLPSRQKK  
ADEPDDISTVV

### **Signal sequence:**

amino acids 1-33

### **Transmembrane domain:**

amino acids 205-220

### **N-glycosylation site.**

amino acids 47-51, 94-98

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 199-203

### **Casein kinase II phosphorylation site.**

amino acids 162-166, 175-179

### **N-myristoylation site.**

amino acids 37-43, 45-51, 110-116

protein database



## FIGURE 29

ACCGAGCCGAGCGGACCGAAGGCGCGCCCGAGATGCAGGTGAGCAAGAGGATGCTGGCGGGG  
GGCGTGAGGAGCATGCCCAGCCCCCTCCTGGCCTGCTGGCAGCCCATCCTCCTGCTGGTGCT  
GGGCTCAGTGCTGTCAGGCTCGGCCACGGGCTGCCCCGCGCTGCGAGTGCTCCGCCCAGG  
ACCGCGCTGTGCTGTGCCACCGCAAGTGCTTTGTGGCAGTCCCCGAGGGCATCCCCACCGAG  
ACGCGCCTGCTGGACCTAGGCAAGAACC GCATCAAAACGCTCAACCAGGACGAGTTCGCCAG  
CTTCCCGCACCTGGAGGAGCTGGAGCTCAACGAGAACATCGTGAGCGCCGTGGAGCCCCGGCG  
CCTTCAACAACCTCTTCAACCTCCGGACGCTGGGTCTCCGCAGCAACCGCCTGAAGCTCATC  
CCGCTAGGCGTCTTCACTGGCCTCAGCAACCTGACCAAGCAGGACATCAGCGAGAACAAGAT  
CGTTATCCTACTGGACTACATGTTTCAGGACCTGTACAACCTCAAGTCACTGGAGGTTGGCG  
ACAATGACCTCGTCTACATCTCTCACC GCGCCTTCAGCGGCCTCAACAGCCTGGAGCAGCTG  
ACGCTGGAGAAATGCAACCTGACCTCCATCCCCACCGAGGCGCTGTCCCACCTGCACGGCCT  
CATCGTCCTGAGGCTCCGGCACCTCAACATCAATGCCATCCGGGACTACTCCTTCAAGAGGC  
TGTACCGACTCAAGGTCTTGGAGATCTCCCACTGGCCCTACTTGGACACCATGACACCCAAC  
TGCCTCTACGGCCTCAACCTGACGTCCCTGTCCATCACACACTGCAATCTGACCGCTGTGCC  
CTACCTGGCCGTCCGCCACCTAGTCTATCTCCGCTTCTCTCAACCTCTCCTACAACCCCATCA  
GCACCATTTAGGGGCTCCATGTTGCATGAGCTGCTCCGGCTGCAGGAGATCCAGCTGGTGGGC  
GGGCAGCTGGCCGTGGTGGAGCCCTATGCCTTCCGCGGCCTCAACTACCTGCGCGTGCTCAA  
TGTCTCTGGCAACCAGCTGACCACACTGGAGGAATCAGTCTTCCACTCGGTGGGCAACCTGG  
AGACACTCATCCTGGACTCCAACCCGCTGGCCTGCGACTGTGCGCTCCTGTGGGTGTTCCGG  
CGCCGCTGGCGGCTCAACTTCAACCGGCAGCAGCCCACGTGCGCCACGCCCCGAGTTTGTCCA  
GGGCAAGGAGTTCAAGGACTTCCCTGATGTGCTACTGCCCAACTACTTCACTGCGCGCGCG  
CCCGCATCCGGGACCGCAAGGCCCAGCAGGTGTTTGTGGACGAGGGCCACACGGTGCAGTTT  
GTGTGCCGGGCCGATGGCGACCCGCGCCCGCCATCCTCTGGCTCTCACCCCGAAAGCACCT  
GGTCTCAGCCAAGAGCAATGGGCGGCTCACAGTCTTCCCTGATGGCACGCTGGAGGTGCGCT  
ACGCCCAGGTACAGGACAACGGCACGTACCTGTGCATCGCGGCCAACGCGGGCGGCAACGAC  
TCCATGCCCCGCCACCTGCATGTGCGCAGCTACTCGCCCCGACTGGCCCCATCAGCCCAACAA  
GACCTTCGCTTTCATCTCCAACCAGCCGGGCGAGGGAGAGGCCAACAGCACCCGCGCCACTG  
TGCCTTTCCCCTTCGACATCAAGACCCTCATCATCGCCACCACCATGGGCTTCATCTCTTTC  
CTGGGCGTCGTCTCTTCTGCCTGGTGCTGCTGTTTCTCTGGAGCCGGGGCAAGGGCAACAC  
AAAGCACAACATCGAGATCGAGTATGTGCCCCGAAAGTCGGACGCAGGCATCAGCTCCGCCG  
ACGCGCCCCGCAAGTTCAACATGAAGATGATATGAGGCGCGGGCGGGGGCAGGGACCCCCG  
GGCGGCCCGGCAGGGGAAGGGGCCTGGTCGCCACCTGCTCACTCTCCAGTCTTCCCACCTC  
CTCCCTACCCTTCTACACACGTTCTCTTTCTCCCTCCCGCCTCCGTCCCCTGCTGCCCCCG  
CCAGCCCTCACCACCTGCCCTCCTTCTACCAGGACCTCAGAAGCCCAGACCTGGGGACCCCA  
CCTACACAGGGGCATTGACAGACTGGAGTTGAAAGCCGACGAACCGACACGCGGCAGAGTCA  
ATAATTCAATAAAAAAGTTACGAACCTTCTCTGTAACCTTGGGTTTCAATAATTATGGATTTT  
TATGAAAACCTTGAAATAATAAAAAGAGAAAAAACTAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 30**

MQVSKRMLAGGVRSMPSPLLACWQPILLVLGSLSGSATGCPPRCECSAQDRAVLCHRKCF  
VAVPEGIPTETRLLDLGNRIKTLNQDEFASFPHLEELNENIVSAVEPGAFFNNLFNLRTL  
GLRSNRLKLIPLGVFTGLSNLTQQDISENKIVILLDYMFDLYNLKSLEVGDNDLVYISHRA  
FSGNLNSLEQLTLEKCNLTSLIPTALSHLHGLIVLRLRHLNINAIIRDYSFKRLYLKVLKLEISH  
WPYLDTMTNCLYGLNLTSLSTHNCNLTAVPYLAVRHLVYLRLNLSYNPISTIEGSMLEL  
LRLQEIQLVGGQLAVVEPYAFRGLNYLRVLNVSGNQLTTLEESVFHSGNLETLILDSNPLA  
CDCRLLWVFRRRWRLNFRNRQQPTCATPEFVQGKEFKDFPDVLLPNYFTCRRARIRDRKAQQV  
FVDEGHTVQFVCRADGDPPIALWLSRKHLSAKSNGRLTVFPDGTLEVRYAQVDNGTYL  
CIAANAGGNDSPAHLHVRSYSPDWPHQPNKTFAFISNQPGEGEANSTRATVPFPFDIKTLI  
IATTMGFISFLGVVLFCLVLLFLWSRGKGNTKHNIIEYVPRKSDAGISSADAPRKFNMKMI

### **Signal sequence:**

amino acids 1-41

### **Transmembrane domain:**

amino acids 556-578

### **N-glycosylation site.**

amino acids 144-148, 202-206, 264-268, 274-278, 293-297, 341-345,  
492-496, 505-509, 526-530, 542-546

### **Casein kinase II phosphorylation site.**

amino acids 49-53, 108-112, 146-150, 300-304, 348-352, 349-353,  
607-611

### **Tyrosine kinase phosphorylation site.**

amino acids 590-598

### **N-myristoylation site.**

amino acids 10-16, 32-38, 37-43, 113-119, 125-131, 137-143,  
262-268, 320-326, 344-350, 359-365, 493-499, 503-509, 605-611

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 32-43

## FIGURE 31

CCCACGCGTCCGCACCTCGGCCCCGGGCTCCGAAGCGGCTCGGGGGCGCCCTTTCGGTCAAC  
ATCGTAGTCCACCCCCTCCCCATCCCCAGCCCCGGGGATTAGGCTCGCCAGCGCCCAGCC  
AGGGAGCCGGCCGGGAAGCGCGATGCGGGGGCCCCAGCCGCCTCGCTCCTGCTCCTGCTCCTGC  
TGTTTCGCTGCTGCTGGGCGCCCCGGCGGGGCCAACCTCTCCCAGGACGACAGCCAGCCCTGG  
ACATCTGATGAAACAGTGGTGGCTGGTGGCACCGTGGTGCTCAAGTGCCAAGTGAAAGATCA  
CGAGGACTCATCCCTGCAATGGTCTAACCTGCTCAGCAGACTCTCTACTTTGGGGAGAAGA  
GAGCCCTTCGAGATAATCGAATTCAGCTGGTTACCTCTACGCCCCACGAGCTCAGCATCAGC  
ATCAGCAATGTGGCCCTGGCAGACGAGGGCGAGTACACCTGCTCAATCTTCACTATGCCTGT  
GCGAACTGCCAAGTCCCTCGTCACTGTGCTAGGAATTCACAGAAGCCCATCATCACTGGTT  
ATAAATCTTCATTACGGGAAAAAGACACAGCCACCCTAACTGTCAGTCTTCTGGGAGCAAG  
CCTGCAGCCCCGGCTCACCTGGAGAAAGGGTGACCAAGAACTCCACGGAGAACCAACCCGCAT  
ACAGGAAGATCCCAATGGTAAACCTTCACTGTCAGCAGCTCGGTGACATTCCAGGTTACCC  
GGGAGGATGATGGGGCGAGCATCGTGTGCTCTGTGAACCATGAATCTCTAAAGGGAGCTGAC  
AGATCCACCTCTCAACGCATTGAAGTTTTATACACACCAACTGCGATGATTAGGCCAGACCC  
TCCCCATCCTCGTGAGGGCCAGAAGCTGTTGCTACACTGTGAGGGTTCGCGGCAATCCAGTCC  
CCCAGCAGTACCTATGGGAGAAGGAGGGCAGTGTGCCACCCCTGAAGATGACCCAGGAGAGT  
GCCCTGATCTTCCCTTTCCTCAACAAGAGTGACAGTGGCACCTACGGCTGCACAGCCACCAG  
CAACATGGGCAGCTACAAGGCCTACTACACCCTCAATGTTAATGACCCCAGTCCGGTGGCCCT  
CCTCCTCCAGCACCTACCACGCCATCATCGGTGGGATCGTGGCTTTCATTGTCTTCTGCTG  
CTCATCATGCTCATCTTCTTGGCCACTACTTGATCCGGCACAAAGGAACCTACCTGACACA  
TGAGGCAAAAGGCTCCGACGATGCTCCAGACGCGGACACGGCCATCATCAATGCAGAAGGCG  
GGCAGTCAGGAGGGGACGACAAGAAGGAATATTTTCATCTAGAGGCGCCTGCCCACTTCTGCTG  
GCCCCCAGGGGGCCCTGTGGGGACTGCTGGGGCCGTCACCAACCCGGACTTGTAACAGAGCAA  
CCGCAGGGCCGCCCCTCCCGCTTGCTCCCCAGCCCACCCACCCCTGTACAGAATGTCTGC  
TTTGGGTGCGGTTTTGTACTCGGTTTGGAATGGGGAGGGAGGAGGGCGGGGGAGGGGAGGG  
TTGCCCTCAGCCCTTTCGCTGGCTTCTCTGCATTTGGGTTATTATTATTTTTGTAAACAATCC  
CAAATCAAATCTGTCTCCAGGCTGGAGAGGCAGGAGCCCTGGGGTGAGAAAAGCAAAAACA  
AACAAAAACA

## **FIGURE 32**

MGAPAASLLLLLLLLFACCWAPGGANLSQDDSQPWTSDET VVAGGTVVLKCQVKDHEDSSLQW  
SNPAQQTLYFGEKRALRDNRIQLVTSTPHELSSISISNVALADEGEYTC SIFTMPVRTAKSLV  
TVLGIPQKPIITGYKSSLREKDTATLNCQSSGSKPAARLTWRKGDQELHGEPTRI QEDPNGK  
TFTVSSSVTFQVTREDDGASIVCSVNHESLKGADRSTSQR IEVLYTPTAMIRPDPPHPREGQ  
KLLHCEGRGNPVPQQYLWEKEGSVPPLKMTQESALIFPFLNKSDSGTYGCTATS NMGSYKA  
YYTLNVNDPSPVPSSSSTYHAIIGGIVAFIVFLLIMLI FLGHYLIRHKGTYLT HEAKGSDD  
APDADTAIINAEGGQSGGDDKKEYFI

### **Signal sequence:**

amino acids 1-20

### **Transmembrane domain:**

amino acids 331-352

### **N-glycosylation site.**

amino acids 25-29, 290-294

### **Casein kinase II phosphorylation site.**

amino acids 27-31, 35-39, 89-93, 141-145, 199-203, 388-392

### **N-myristoylation site.**

amino acids 2-8, 23-29, 156-162, 218-224, 295-301, 298-304,  
306-310, 334-340, 360-364, 385-389, 386-390

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 7-18

## FIGURE 33

GGGGGTTAGGGAGGAAGGAATCCACCCCCACCCCCCAAACCCTTTTCTTCTCCTTTCCTGG  
 CTTCGGACATTGGAGCACTAAATGAACTTGAATTGTGTCTGTGGCGAGCAGGATGGTCGCTG  
 TTACTTTGTGATGAGATCGGGGATGAATTGCTCGCTTTAAAAATGCTGCTTTGGATTCTGTT  
 GCTGGAGACGTCTCTTTGTTTTGCCGCTGGAAACGTTACAGGGGACGTTTGCAAAGAGAAGA  
 TCTGTTTCTGCAATGAGATAGAAGGGGACCTACACGTAGACTGTGAAAAAAGGGCTTCACA  
 AGTCTGCAGCGTTTCACTGCCCCGACTTCCCAGTTTTTACCATTTATTTCTGCATGGCAATTC  
 CCTCACTCGACTTTTCCCTAATGAGTTTCGCTAACTTTTATAATGCGGTTAGTTTGCACATGG  
 AAAACAATGGCTTGCATGAAATCGTTCCGGGGGCTTTTCTGGGGCTGCAGCTGGTGAAAAGG  
 CTGCACATCAACAACAAGATCAAGTCTTTTCGAAAGCAGACTTTTCTGGGGCTGGACGA  
 TCTGGAATATCTCCAGGCTGATTTTAAATTTATTACGAGATATAGACCCGGGGGCTTCCAGG  
 ACTTGAACAAGCTGGAGGTGCTCATTTTAAATGACAATCTCATCAGCACCTTACCTGCCAAC  
 GTGTTCCAGTATGTGCCCATCACCCACCTCGACCTCCGGGGTAACAGGCTGAAAACGCTGCC  
 CTATGAGGAGGTCTTGGAGCAAATCCCTGGTATTGCGGAGATCCTGCTAGAGGATAACCTTT  
 GGGACTGCACCTGTGATCTGCTCTCCCTGAAAGAATGGCTGGAAAACATTCCCAAGAATGCC  
 CTGATCGGCCGAGTGGTCTGCGAAGCCCCCACCAGACTGCAGGGTAAAGACCTCAATGAAAC  
 CACCGAACAGGACTTGTGTCTTTGAAAAACCGAGTGGATTCTAGTCTCCCGGCGCCCCCTG  
 CCAAGAAGAGACCTTTGCTCCTGGACCCCTGCCAACTCCTTTCAAGACAAATGGGCAAGAG  
 GATCATGCCACACCAGGGTCTGCTCCAAACGGAGGTACAAAGATCCCAGGCAACTGGCAGAT  
 CAAAATCAGACCCACAGCAGCGATAGCGACGGGTAGCTCCAGGAACAAACCCTTAGCTAACA  
 GTTTACCCTGCCCTGGGGGCTGCAGCTGCGACCACATCCCAGGGTTCGGGTTTAAAGATGAAC  
 TGCAACAACAGGAACGTGAGCAGCTTGGCTGATTTGAAGCCCAAGCTCTCTAACGTGCAGGA  
 GCTTTTCTACGAGATAACAAGATCCACAGCATCCGAAAATCGCACTTTGTGGATTACAAGA  
 ACCTCATTCTGTTGGATCTGGGCAACAATAACATCGCTACTGTAGAGAACAACACTTTCAAG  
 AACCTTTTGGACCTCAGGTGGCTATACATGGATAGCAATTACCTGGACACGCTGTCCCGGGA  
 GAAATTCGCGGGGCTGCAAAACCTAGAGTACCTGAACGTGGAGTACAACGCTATCCAGCTCA  
 TCCTCCCGGGCACTTTCAATGCCATGCCCAAACCTGAGGATCCTCATTCTCAACAACAACCTG  
 CTGAGGTCCCTGCCGTGTGGACGTGTTTCGCTGGGGTCTCGCTCTCTAAACTCAGCCTGCACAA  
 CAATTACTTTCATGTACCTCCCGGTGGCAGGGGTGCTGGACCAGTTAACCTCCATCATCCAGA  
 TAGACCTCCACGGAAACCCCTGGGAGTGCTCCTGCACAATTGTGCCCTTCAAGCAGTGGGCA  
 GAACGCTTGGGTTCCGAAGTGCTGATGAGCGACCTCAAGTGTGAGACGCCGGTGAACCTTCTT  
 TAGAAAGGATTTTATGCTCCTCTCCAATGACGAGATCTGCCCTCAGCTGTACGCTAGGATCT  
 CGCCACGTTAACTTCGCACAGTAAAAACAGCACTGGGTTGGCGGAGACCGGGACGCACTCC  
 AACTCCTACCTAGACACCAGCAGGGTGTCCATCTCGGTGTTGGTCCCGGGACTGCTGCTGGT  
 GTTTGTACCTCCGCCTTCACCGTGGTGGGCATGCTCGTGTATCCTGAGGAACCGAAAGC  
 GGTCCAAGAGACGAGATGCCAACTCCTCCGCGTCCGAGATTAATTCCCTACAGACAGTCTGT  
 GACTCTTCTACTGGCACAATGGGCCTTACAACGCAGATGGGGCCACAGAGTGTATGACTG  
 TGGCTCTCACTCGCTCTCAGACTAAAGACCCCAACCCCAATAGGGGAGGGCAGAGGGAAGGCG  
 ATACATCCTTCCCCACCGCAGGCACCCCGGGGGCTGGAGGGGCGTGTACCCAAATCCCCGCG  
 CCATCAGCCTGGATGGGCATAAGTAGATAAAATAACTGTGAGCTCGCACAACCGAAAGGGCCT  
 GACCCCTTACTTAGCTCCCTCCTTGAAACAAAGAGCAGACTGTGGAGAGCTGGGAGAGCGCA  
 GCCAGCTCGCTCTTTGCTGAGAGCCCCCTTTTGACAGAAAGCCCAGCACGACCTGTCTGGAAG  
 AACTGACAGTGCCCTCGCCCTCGGCCCGGGGCTGTGGGGTTGGATGCCGCGGTTCTATAC  
 ATATATACATATATCCACATCTATATAGAGAGATAGATATCTATTTTTCCCTGTGGATTAG  
 CCGCGTGATGGCTCCCTGTTGGCTACGCAGGGATGGGCAGTTGCACGAAGGCATGAATGTAT  
 TGTAATAAGTAACTTTGACTTCTGAC

## **FIGURE 34**

MLLWILLLETSLCFAAGNVTGDVCKEIKCSCNEIEGDLHVDCEKKGFTSLQRFTAPTSQFYH  
LFLHGNSLTRLPNEFANFYNAVSLHMENGLHEIVPGAFLGLQLVKRLHINNNKIKSFRKQ  
TFLGLDDLEYLQADFNLLRDIDPGAQDLNKLEVLILNDNLISTLPANVFQYVPITHLDLRG  
NRLKTLPEYEEVLEQIPGIAEILLEDNPWDCTCDLLSLKEWLENIPKNALIGRVVCEAPTRLQ  
GKDLNETTEQDLCPLKNRVDSSLPAPPAQEETFAPGPLPTPFKTNGQEDHATPGSAPNGGTK  
IPGNWQIKIRPTAAIATGSSRNKPLANSRPCPGGCSCDHIPGSGLKMNCCNNRVSSSLADLKP  
KLSNVQELFLRDNKHISIRKSHFVDYKNLILLDLGNNNIATVENNTFKNLLDLRWLYMDSNY  
LDTLSREKFAGLQNLLEYLNVEYNAIQILPGTFNAMPKLRIILNNNLLRSLPVDVFAGVSL  
SKLSLHNNYFMYLPVAGVLDQLTSIIQIDLHGPNWECSTIVPFKQWAERLGSEVLMSDLKC  
ETPVNFFRKDFMLLSNDEICPQLYARISPTLTSHSKNSTGLAETGTHSNSYLDTSRVSISVL  
VPGLLLVFTVSAFTTVGMLVFILNRNRKRSKRRDANSSASEINSLQTVCDSSYWHNGPYNADG  
AHRVYDCGSHSLSD

### **Signal sequence:**

amino acids 1-15

### **Transmembrane domain:**

amino acids 618-638

### **N-glycosylation site.**

amino acids 18-22, 253-257, 363-367, 416-420, 595-599, 655-659

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 122-126, 646-650

### **Casein kinase II phosphorylation site.**

amino acids 30-34, 180-184, 222-226, 256-260, 366-370, 573-577,  
608-612, 657-661, 666-670, 693-697

### **N-myristoylation site.**

amino acids 17-23, 67-73, 100-106, 302-308, 328-334, 343-349,  
354-360, 465-471, 493-499, 598-604, 603-609

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 337-348

## FIGURE 35

AGTCGACTGCGTCCCCTGTACCCGGCGCCAGCTGTGTTCTTGACCCCAGAATAACTCAGGGC  
TGCACCGGGCCTGGCAGCGCTCCGCACACATTTCTGTGCGGGCCTAAGGGAAACTGTTGGC  
CGCTGGGCCCCGCGGGGGGATTCTTGGCAGTTGGGGGGTCCGTGCGGAGCGAGGGCGGAGGGG  
AAGGGAGGGGGAACCGGGTTGGGGAAGCCAGCTGTAGAGGGCGGTGACCGCGCTCCAGACAC  
AGCTCTGCGTCCCTCGAGCGGGACAGATCCAAGTTGGGAGCAGCTCTGCGTGCGGGGCCTCAG  
AGAATGAGGCGCGGCGTTTCGCCCTGTGCCTCCTCTGGCAGGCGCTCTGGCCCCGGGCGGGCGG  
CGGCGAACACCCCACTGCCGACCGTGCTGGCTGCTCGGCCTCGGGGGCCTGCTACAGCCTGC  
ACCACGCTACCATGAAGCGGCAGGCGGCCGAGGAGGCCTGCATCCTGCGAGGTGGGGCGCTC  
AGCACCGTGCGTGCGGGCGCCGAGCTGCGCGCTGTGCTCGCGCTCCTGCGGGCAGGCCCAGG  
GCCCCGAGGGGGCTCCAAAGACCTGCTGTTCTGGGTGCGACTGGAGCGCAGGCGTTCCCACT  
GCACCCTGGAGAACGAGCCTTTGCGGGGTTTCTCCTGGCTGTCTCCGACCCCGGCGGTCTC  
GAAAGCGACACGCTGCAGTGGGTGGAGGAGCCCCAACGCTCCTGCACCGCGCGGAGATGCGC  
GGTACTCCAGGCCACCGGTGGGGTGCAGCCCGCAGGCTGGAAGGAGATGCGATGCCACCTGC  
GCGCCAACGGCTACCTGTGCAAGTACCAGTTTGAGGTCTTGTGTCCTGCGCCGCGCCCCGGG  
GCCGCCTCTAACTTGAGCTATCGCGCGCCCTTCCAGCTGCACAGCGCCGCTCTGGACTTCAG  
TCCACCTGGGACCGAGGTGAGTGCGCTCTGCCGGGGACAGCTCCCGATCTCAGTTACTTGCA  
TCGCGGACGAAATCGGCGCTCGCTGGGACAAACTCTCGGGCGATGTGTTGTGTCCTGCCCC  
GGGAGGTACCTCCGTGCTGGCAAATGCGCAGAGCTCCCTAACTGCCTAGACGACTTGGGAGG  
CTTTGCCTGCGAATGTGCTACGGGCTTCGAGCTGGGGAAGGACGGCCGCTCTTGTGTGACCA  
GTGGGGAAGGACAGCCGACCTTGGGGGGACCGGGGTGCCACCAGGCGCCCGCGGCCACT  
GCAACCAGCCCCGTGCCGAGAGAACATGGCCAATCAGGGTCGACGAGAAGCTGGGAGAGAC  
ACCACTTGTCCCTGAACAAGACAATTTCAGTAACATCTATTCTGAGATTCCTCGATGGGGAT  
CACAGAGCACGATGTCTACCCTTCAAATGTCCCTTCAAGCCGAGTCAAAGGCCACTATCACC  
CCATCAGGGAGCGTGATTTCCAAGTTTAAATTCTACGACTTCCTCTGCCACTCCTCAGGCTTT  
CGACTCCTCCTCTGCCGTGGTCTTCATATTTGTGAGCACAGCAGTAGTAGTGTGGTGATCT  
TGACCATGACAGTACTGGGGCTTGTCAAGCTCTGCTTTCACGAAAGCCCCCTCTTCCCAGCCA  
AGGAAGGAGTCTATGGGCCCCGCGGGCCTGGAGAGTGATCCTGAGCCCGCTGCTTTGGGCTC  
CAGTTCTGCACATTGCACAAACAATGGGGTGAAAGTCGGGGACTGTGATCTGCGGGACAGAG  
CAGAGGGTGCCTTGCTGGCGGAGTCCCTCCTTGGCTCTAGTGATGCATAGGGGAAACAGGGGA  
CATGGGCACTCCTGTGAACAGTTTTTCACTTTTGATGAAACGGGGAACCAAGAGGAACTTAC  
TTGTGTAAC TGACAATTTCTGCAGAAATCCCCCTTCTCTAAATTCCCTTTACTCCACTGAG  
GAGCTAAATCAGAACTGCACACTCCTTCCCTGATGATAGAGGAAGTGGAAGTGCCTTTAGGA  
TGGTGATACTGGGGGACCGGGTAGTGCTGGGGAGAGATATTTTCTTATGTTTATTCGGAGAA  
TTTGGAGAAGTGATTGAACTTTTCAAGACATTGGAAACAAATAGAACACAATATAATTTACA  
TTAAAAAATAATTTCTACCAAATGGAAAGGAAATGTTCTATGTTGTTTCAGGCTAGGAGTAT  
ATTGGTTCGAAATCCCAGGGAAAAAATAAAAAATAAAAAATTAAAGGATTGTTGAT

## **FIGURE 36**

MRPAFALCLLWQALWPGPGGGEHPTADRAGCSASGACYSLHHATMKRQAEEACILRGGALS  
TVRAGAE LR AVLALLRAGPGPGGSKDLLFWVALERRRSHCTLENEPLRGFSWLSSDPGGLE  
SDTLQWVEEPQRSCTARRCAVLQATGGVEPAGWKEMRCHLRANGYLCKYQFEVLC PAPRPGA  
ASNLSYRAPFQLHSAALDFSPPGTEVSALCRGQLPISVTCTIADEIGARWDKLSGDVLCPCPG  
RYLRAGKCAELPNCLDDLGGFACECATGFELGKDGRSCVTS GEGQPTLGGTGVPTRRPPATA  
TSPVPQRTWPPIRVDEKLGETPLVPEQDNSVTISIPEIPRWGSQSTMSTLQMSLQAESKATITP  
SGSVISKFNSTTSSATPQAFDSSSAVVFI FVSTAVVVLVILTM TVLGLVKLCFHESPSSQPR  
KESMGPPGLESDPEPAALGSSSAHCTNNGVKVGDCDLRDRAEGALLAESPLGSSDA

### **Signal sequence:**

amino acids 1-16

### **Transmembrane domain:**

amino acids 399-418

### **N-glycosylation site.**

amino acids 189-193, 381-385

### **Glycosaminoglycan attachment site.**

amino acids 289-293

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 98-102, 434-438

### **Casein kinase II phosphorylation site.**

amino acids 275-279, 288-292, 342-346, 445-449

### **N-myristoylation site.**

amino acids 30-36, 35-41, 58-64, 59-65, 121-127, 151-157,  
185-191, 209-215, 267-273, 350-356, 374-380, 453-459, 463-469,  
477-483

### **Aspartic acid and asparagine hydroxylation site.**

amino acids 262-274



## FIGURE 37

CGGACGCGTGGGATTTCAGCAGTGGCCTGTGGCTGCCAGAGCAGCTCCTCAGGGGAAACTAAG  
CGTCGAGTCAGACGGCACCATAATCGCCTTTAAAAGTGCCTCCGCCCTGCCGGCCGCGTATC  
CCCCGGCTACCTGGGCGCGCCCGCGGCGGTGCGCGCGTGAGAGGGAGCGCGCGGGCAGCCGA  
GCGCCGGTGTGAGCCAGCGCTGCTGCCAGTGTGAGCGGCGGTGTGAGCGCGGTGGGTGCGGA  
GGGGCGTGTGTGCCGGCGCGCGCGCCGTGGGGTGCAAACCCCGAGCGTCTACGCTGCCATGA  
GGGGCGCGAACGCTGGGCGCCACTCTGCCTGCTGCTGGCTGCCGCCACCCAGCTCTCGCGG  
CAGCAGTCCCCAGAGAGACCTGTTTTTACATGTGGTGGCATTCTTACTGGAGAGTCTGGATT  
TATTGGCAGTGAAGGTTTTCTGGAGTGTACCTCCAAATAGCAAATGTACTTGGAATCA  
CAGTTCCTGAAGGAAAAGTAGTCGTTCTCAATTTCCGATTCATAGACCTCGAGAGTGACAAC  
CTGTGCCGCTATGACTTTGTGGATGTGTACAATGGCCATGCCAATGGCCAGCGCATTTGGCCG  
CTTCTGTGGCACTTTCCGGCCTGGAGCCCTTGTGTCCAGTGGCAACAAGATGATGGTGCAGA  
TGATTTCTGATGCCAACACAGCTGGCAATGGCTTCATGGCCATGTTCTCCGCTGCTGAACCA  
AACGAAAGAGGGGATCAGTATTGTGGAGGACTCCTTGACAGACCTTCCGGCTCTTTTAAAAC  
CCCCAACTGGCCAGACCGGGATTACCCTGCAGGAGTCACTTGTGTGTGGCACATTGTAGCCC  
CAAAGAATCAGCTTATAGAATTAAAGTTTGAGAAGTTTGATGTGGAGCGAGATAACTACTGC  
CGATATGATTATGTGGCTGTGTTTAAATGGCGGGGAAGTCAACGATGCTAGAAGAATTGAAA  
GTATTGTGGTGATAGTCCACCTGCGCCAATTGTGTCTGAGAGAAATGAACTTCTTATTCAGT  
TTTTATCAGACTTAAGTTTAACTGCAGATGGGTTTATTGGTCACTACATATTCAGGCCAAAA  
AACTGCCTACAACACAGAACAGCCTGTCACCACCACATTCCCTGTAACCACGGGTTTAAA  
ACCCACCGTGGCCTTGTGTCAACAAAAGTGTAGACGGACGGGGACTCTGGAGGGCAATTATT  
GTTCAAGTGACTTTGTATTAGCCGGCACTGTTATCACAAACCATCACTCGCGATGGGAGTTTG  
CACGCCACAGTCTCGATCATCAACATCTACAAAGAGGGAAATTTGGCGATTACAGCAGGCGGG  
CAAGAACATGAGTGCCAGGCTGACTGTCGTCTGCAAGCAGTGCCCTCTCCTCAGAAGAGGTC  
TAAATTACATTATTATGGGCCAAGTAGGTGAAGATGGGCGAGGCAAAATCATGCCAAACAGC  
TTTATCATGATGTTCAAGACCAAGAATCAGAAGCTCCTGGATGCCTTAAAAAATAAGCAATG  
TTAAACAGTGAAGTGTGTCCATTTAAGCTGTATTCTGCCATTGCCTTTGAAAGATCTATGTTT  
TCTCAGTAGAAAAAAAATACTTATAAAATTACATATTCTGAAAGAGGATTCCGAAAGATGG  
GACTGGTTGACTCTTCACATGATGGAGGTATGAGGCCTCCGAGATAGCTGAGGGAAGTTCTT  
TGCCTGCTGTCAGAGGAGCAGCTATCTGATTGGAAACCTGCCGACTTAGTGCGGTGATAGGA  
AGCTAAAAGTGTCAAGCGTTGACAGCTTGGAAGCGTTTATTTATACATCTCTGTAAAAGGAT  
ATTTTAGAATTGAGTTGTGTGAAGATGTCAAAAAAGATTTTAGAAGTGCAATATTTATAGT  
GTTATTTGTTTCACCTTCAAGCCTTTGCCCTGAGGTGTTACAATCTTGTCTTGCGTTTTCTA  
AATCAATGCTTAATAAAATATTTTAAAGGAAAAAAAAAAAAA

“F3” chimera

## **FIGURE 38**

MRGANAWAPLCLLLAAATQLSRQQSPERPVFCTCGGILTGESGFIGSEGFPGVYPPNSKCTWK  
ITVPEGKVVLNFRFIDLESDNLCRYDFVDVYNGHANGQRIGRFCGTFRPGALVSSGNKMMV  
QMISDANTAGNGFMAMFSAAEPNERGDQYCGGLLDRPSGSFKTPNWPDRDYPAGVTCVWHIV  
APKNQLIELKFEKFDVERDNYCRYDYVAVFNGGEVNDARRIGKYCGDSPAPPIVSENERELLI  
QFLSDLSLTADGFIGHYIFRPKKLPPTTTEQPVTTTTFPVTTLGLKPTVALCQQKCRRTGTLEGN  
YCSSDFVLGTVITTITRDGSLHATVSIINIYKEGNLAIQQAGKNMSARLTVVCKQCPLLRR  
GLNYIIMGQVGEDGRGKIMPNSFIMMFKTKNQKLLDALKNKQC

### **Signal sequence:**

amino acids 1-23

### **N-glycosylation site.**

amino acids 355-359

### **Casein kinase II phosphorylation site.**

amino acids 64-68, 142-146, 274-278

### **Tyrosine kinase phosphorylation site.**

amino acids 199-208

### **N-myristoylation site.**

amino acids 34-40, 35-41, 100-106, 113-119, 218-224, 289-295,  
305-311, 309-315, 320-326, 330-336

### **Cell attachment sequence.**

amino acids 149-152

protein database

# POPE JOHN

CGGACGCGTGGGCGGACGCGTGGGCGGCCACGGCGCCCGCGGGCTGGGGCGGTTCGCTTCTT  
CCTTCTCCGTGGCCTACGAGGGTCCCCAGCCTGGGTAAAGATGGCCCCATGGCCCCCGAAGG  
GCCTAGTCCCAGCTGTGCTCTGGGGCCTCAGCCTCTTCCTCAACCTCCCAGGACCTATCTGG  
CTCCAGCCCTCTCCACCTCCCCAGTCTTCTCCCCCGCCTCAGCCCCATCCGTGTTCATACCTG  
CCGGGGACTGGTTGACAGCTTTAACAAGGGCCTGGAGAGAACCATCCGGGACAACCTTTGGAG  
GTGGAACAACACTGCCTGGGAGGAAGAGAATTTGTCCAAATACAAAGACAGTGAGACCCGCCTG  
GTAGAGGTGCTGGAGGGTGTGTGCAGCAAGTCAGACTTCGAGTGCCACCGCCTGCTGGAGCT  
GAGTGAGGAGCTGGTGGAGAGCTGGTGGTTTCACAAGCAGCAGGAGGCCCCGGACCTCTTCC  
AGTGGCTGTGCTCAGATTCCCTGAAGCTCTGCTGCCCCGCAGGCACCTTCGGGCCCTCCTGC  
CTTCCCTGTCTGGGGGAACAGAGAGGCCCTGCGGTGGCTACGGGCAGTGTGAAGGAGAAGG  
GACACGAGGGGGCAGCGGGCACTGTGACTGCCAAGCCGGCTACGGGGGTGAGGCCTGTGGCC  
AGTGTGGCCTTGGCTACTTTGAGGCAGAACGCAACGCCAGCCATCTGGTATGTTTCGGCTTGT  
TTTGGCCCCCTGTGCCCGATGCTCAGGACCTGAGGAATCAAACCTGTTTGCAATGCAAGAAGGG  
CTGGGGCCCTGCATCACCTCAAGTGTGTAGACATTGATGAGTGTGGCACAGAGGGAGCCAACT  
GTGGAGCTGACCAATTCTGCGTGAACACTGAGGGCTCCTATGAGTGCCGAGACTGTGCCAAG  
GCCTGCCTAGGCTGCATGGGGGCAGGGCCAGGTGCTGTAAGAAGTGTAGCCCTGGCTATCA  
GCAGGTGGGCTCCAAGTGTCTCGATGTGGATGAGTGTGAGACAGAGGTGTGTCCGGGAGAGA  
ACAAGCAGTGTGAAAACACCGAGGGCGGTTATCGCTGCATCTGTGCCGAGGGCTACAAGCAG  
ATGGAAGGCATCTGTGTGAAGGAGCAGATCCCAGAGTCAGCAGGCTTCTTCTCAGAGATGAC  
AGAAGACGAGTTGGTGGTGTGCTGCAGCAGATGTTCTTTGGCATCATCATCTGTGCACTGGCCA  
CGCTGGCTGCTAAGGGCGACTTGGTGTTACCGCCATCTTCATTGGGGCTGTGGCGGCCATG  
ACTGGCTACTGGTTGTGAGAGCGCAGTGACCGTGTGCTGGAGGGCTTCATCAAGGGCAGATA  
ATCGCGGCCACCACCTGTAGGACCTCCTCCCACCCACGCTGCCCCCAGAGCTTGGGCTGCCC  
TCCTGCTGGACACTCAGGACAGCTTGGTTTATTTTTGAGAGTGGGGTAAGCACCCCTACCTG  
CCTTACAGAGCAGCCCAGGTACCCAGGCCCGGGCAGACAAGGCCCTGGGGTAAAAAGTAGC  
CCTGAAGGTGGATAACCATGAGCTCTTCACCTGGCGGGGACTGGCAGGCTTCACAATGTGTGA  
ATTTCAAAAGTTTTTTCCTTAATGGTGGCTGCTAGAGCTTTGGCCCCCTGCTTAGGATTAGGTG  
GTCCTCACAGGGGTGGGGCCATCACAGCTCCCTCCTGCCAGCTGCATGCTGCCAGTTCCTGT  
TCTGTGTTACCCACATCCCCACACCCCATTGCCACTTATTTATTTCATCTCAGGAAATAAAGA  
AAGGTCTTGGAAGTTAAAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 40**

MAPWPPKGLVPAVLWGLSLFLNLP GPIWLQPSPPPQSSPPPQPHPCHTCRGLVDSFNKGLER  
TIRDNFGGGNTAWEENLSKYKDSETRLVEVLEGVCSKSDFECHRLLLELSEELVESWWFHKQ  
QEAPDLFQWLCSDSLKLCCPAGTFGPSCLPCPGGTERPCGGYGQCEGEGTRGGSGHCDCQAG  
YGGEACGQCGLGYFEARNASHLVCSACFGPCARCSGPESNCLQCKKGWALHHLKCVDIDE  
CGTEGANCGADQFCVNTESGYECRDCAKACLGCMGAGPGRCKKCSPGYQQVGSKCLDVDECE  
TEVCPGENKQCENTEGGYRCICAEGYKQMEGICVKEQIPESAGFFSEMTEDLVVLQQMFFG  
IIICALATLAAKGDLVFTAIFIGAVAAMTGYWLSERSDRVLEGFIKGR

### **Signal sequence:**

amino acids 1-29

### **Transmembrane domain:**

amino acids 372-395

### **N-glycosylation site.**

amino acids 79-83, 205-209

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 290-294

### **Casein kinase II phosphorylation site.**

amino acids 63-67, 73-77, 99-103, 101-105, 222-226, 359-263

### **N-myristoylation site.**

amino acids 8-14, 51-57, 59-65, 69-75, 70-76, 167-173, 173-179,  
177-183, 188-194, 250-256, 253-259, 267-273, 280-286, 283-289,  
326-332, 372-378, 395-401

### **Aspartic acid and asparagine hydroxylation site.**

amino acids 321-333

### **EGF-like domain cysteine pattern signature.**

amino acids 181-193

## FIGURE 41

TGAGACCCTCCTGCAGCCTTCTCAAGGGACAGCCCCACTCTGCCTCTTGCTCCTCCAGGGCA  
GCACCATGCAGCCCCCTGTGGCTCTGCTGGGCACTCTGGGTGTTGCCCCCTGGCCAGCCCCGGG  
GCCGCCCTGACCGGGGAGCAGCTCCTGGGCAGCCTGCTGCGGCAGCTGCAGCTCAAAGAGGT  
GCCCCACCCTGGACAGGGCCGACATGGAGGAGCTGGTCATCCCCACCCACGTGAGGGGCCAGT  
ACGTGGCCCTGCTGCAGCGCAGCCACGGGGACCGCTCCCGCGGAAAGAGGTTTCAGCCAGAGC  
TTCCGAGAGGTGGCCGGCAGGTTCTTGGCGTTGGAGGCCAGCACACACCTGCTGGTGTTCGG  
CATGGAGCAGCGGCTGCCGCCCAACAGCGAGCTGGTGCAGGCCGTGCTGCGGCTCTTCCAGG  
AGCCGGTCCCCAAGGCCGCGCTGCACAGGCACGGGCGGCTGTCCCCGCGCAGCGCCCCGGGCC  
CGGGTGACCGTCGAGTGGCTGCGCGTCCGCGACGACGGCTCCAACCGCACCTCCCTCATCGA  
CTCCAGGCTGGTGTCCGTCCACGAGAGCGGCTGGAAGGCCTTCGACGTGACCGAGGCCGTGA  
ACTTCTGGCAGCAGCTGAGCCGGCCCCGGCAGCCGCTGCTGCTACAGGTGTCGGTGCAGAGG  
GAGCATCTGGGCCCCGCTGGCGTCCGGCGCCACAAGCTGGTCCGCTTTGCCTCGCAGGGGGC  
GCCAGCCGGGCTTGGGGAGCCCCAGCTGGAGCTGCACACCCTGGACCTTGGGGACTATGGAG  
CTCAGGGCGACTGTGACCCTGAAGCACCATGACCGAGGGCACCCGCTGCTGCCGCCAGGAG  
ATGTACATTGACCTGCAGGGGATGAAGTGGGCGGAGAACTGGGTGCTGGAGCCCCCGGGCTT  
CCTGGCTTATGAGTGTGTGGGCACCTGCCGGCAGCCCCCGGAGGCCCTGGCCTTCAAGTGGC  
CGTTTCTGGGGCCTCGACAGTGCATCGCCTCGGAGACTGACTCGCTGCCCATGATCGTCAGC  
ATCAAGGAGGGAGGCAGGACCAGGCCCCAGGTGGTCAGCCTGCCCAACATGAGGGTGCAGAA  
GTGCAGCTGTGCCTCGGATGGTGCCTCGTGCCAAGGAGGCTCCAGCCATAGGGCGCCTAGTG  
TAGCCATCGAGGGACTTGACTTGTGTGTGTTTCTGAAGTGTTTCGAGGGTACCAGGAGAGCTG  
GCGATGACTGAACTGCTGATGGACAAATGCTCTGTGCTCTCTAGTGAGCCCTGAATTTGCTT  
CCTCTGACAAGTTACCTCACCTAATTTTTTGCTTCTCAGGAATGAGAATCTTTGGCCACTGGA  
GAGCCCTTGCTCAGTTTTCTCTATTCTTATTATTCACTGCACTATATTCTAAGCACTTACAT  
GTGGAGATACTGTAACCTGAGGGCAGAAAGCCCCANTGTGTCATTGTTTACTTGTCTGTAC  
TGGATCTGGGCTAAAGTCCTCCACCACCACTCTGGACCTAAGACCTGGGGTTAAGTGTGGGT  
TGTGCATCCCCAATCCAGATAATAAAGACTTTGTAAAACATGAATAAAACACATTTTATTCT  
AAAA

## FIGURE 42

MQPLWLCWALWVLPLASPGAALTGEQLLGSLLRQLQLKEVPTLDRADMEELVIPTHVRAQYV  
ALLQRSHGDRSRGKRFSQSFRVAGRFLALEASTHLLVFGMEQRLPPNSELVQAVLRLFQEP  
VPKAALHRHGRSLSPRSARARVTVEWLRVRDDGSNRTSLIDSRLVSVHESGWKAFDVTEAVNF  
WQQLSRPRQPLLLQVSVQREHLGPLASGAHKLVRFASQGAPAGLGEPQLELHTLDLGDYGAQ  
GDCDPEAPMTEGTRCCRQEMYIDLQGMKWAENWVLEPPGFLAYECVGTCRQPPEALAFKWPF  
LGPRQCIASETDSLPMIVSIKEGGRTRPQVVSLPNMRVQKSCASDGALVPRRLQP

### **Signal sequence:**

amino acids 1-18

### **N-glycosylation site.**

amino acids 158-162

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 76-80

### **Casein kinase II phosphorylation site.**

amino acids 68-72, 81-85, 161-165, 169-173, 319-323, 329-333

### **N-myristoylation site.**

amino acids 19-25, 156-162, 225-231, 260-266, 274-280

### **Amidation site.**

amino acids 74-78

### **TGF-beta family signature.**

amino acids 282-298

## FIGURE 43

GTCTGTTCCCAGGAGTCCTTCGGCGGCTGTTGTGTGTCAGTGGCCTGATCGCGATGGGGACAAA  
GGCGCAAGTCGAGAGGAACTGTTGTGCCTCTTCATATTGGCGATCCTGTTGTGCTCCCTGG  
CATTGGGCAGTGTTACAGTGCACCTTCTGAACCTGAAGTCAGAATTCCTGAGAATAATCCT  
GTGAAGTTGTCCTGTGCCTACTCGGGCTTTTCTTCTCCCCGTGTGGAGTGGAAGTTTGACCA  
AGGAGACACCACCAGACTCGTTTGCTATAATAACAAGATCACAGCTTCCTATGAGGACCGGG  
TGACCTTCTTGCCAACTGGTATCACCTTCAAGTCCGTGACACGGGAAGACACTGGGACATAC  
ACTTGTATGGTCTCTGAGGAAGGCGGCAACAGCTATGGGGAGGTCAAGGTCAAGCTCATCGT  
GCTTGTGCCTCCATCCAAGCCTACAGTTAACATCCCCTCCTCTGCCACCATTGGGAACCGGG  
CAGTGCTGACATGCTCAGAACAAGATGGTTCCCCACCTTCTGAATACACCTGGTTCAAAGAT  
GGGATAGTGATGCCTACGAATCCCAAAGCACCCGTGCCTTCAGCAACTCTTCCTATGTCCT  
GAATCCCACAACAGGAGAGCTGGTCTTTGATCCCCTGTCAGCCTCTGATACTGGAGAATACA  
GCTGTGAGGCACGGAATGGGTATGGGACACCCATGACTTCAAATGCTGTGCGCATGGAAGCT  
GTGGAGCGGAATGTGGGGGTGTCGTGGCAGCCGTCTTGTAAACCCTGATTCTCCTGGGAAT  
CTTGGTTTTTTGGCATCTGGTTTGCCTATAGCCGAGGCCACTTTGACAGAACAAAGAAAGGGA  
CTTCGAGTAAGAAGGTGATTTACAGCCAGCCTAGTGCCCGAAGTGAAGGAGAATTCAAACAG  
ACCTCGTCATTCCTGGTGTGAGCCTGGTTCGGCTCACCGCCTATCATCTGCATTTGCCTTACT  
CAGGTGCTACCGGACTCTGGCCCCCTGATGTCTGTAGTTTCACAGGATGCCTTATTTGTCTTC  
TACACCCACACAGGGCCCCCTACTTCTTCGGATGTGTTTTTAATAATGTCAGCTATGTGCCCC  
ATCCTCCTTCATGCCCTCCCTCCCTTTCCTACCACTGCTGAGTGGCCTGGAACCTTGTTTAAA  
GTGTTTATTCCCCATTTCTTTGAGGGATCAGGAAGGAATCCTGGGTATGCCATTGACTTCCC  
TTCTAAGTAGACAGCAAAAATGGCGGGGGTTCGAGGAATCTGCACTCAACTGCCCACCTGGC  
TGGCAGGGATCTTTGAATAGGTATCTTGAGCTTGTTCTGGGCTCTTTCCTTGTGTACTGAC  
GACCAGGGCCAGCTGTTCTAGAGCGGGAATTAGAGGCTAGAGCGGCTGAAATGGTTGTTTGG  
TGATGACACTGGGGTCCTTCCATCTCTGGGGCCCACTCTCTTCTGTCTTCCCATGGGAAGTG  
CCACTGGGATCCCTCTGCCCTGTCTCCTGAATACAAGCTGACTGACATTGACTGTGTCTGT  
GGAAAATGGGAGCTCTTGTTGTGGAGAGCATAGTAAATTTTCAGAGAAGTTGAAGCCAAAAG  
GATTTAAAACCGCTGCTCTAAAGAAAAGAAAAGTTGGAGGCTGGGCGCAGTGGCTCACGCCTG  
TAATCCCAGAGGCTGAGGCAGGCGGATCACCTGAGGTCGGGAGTTCGGGATCAGCCTGACCA  
ACATGGAGAAACCTACTGGAAATACAAAGTTAGCCAGGCATGGTGGTGCATGCCTGTAGTC  
CCAGCTGCTCAGGAGCCTGGCAACAAGAGCAAACTCCAGCTCAAAAAAAAAAAAAAAAAA

## **FIGURE 44**

MGTKAQVERKLLCLFILAILLCSLALGSVTVHSSEPEVRIPENNPVKLSCAYSGFSSPRVEW  
KFDQGD TTRLVCYNNKITASYEDRVTF LPTGITFKSVTREDTGTYTCMVSEEGGNSYGEVKV  
KLIVLVPPSKPTVNI PSSATIGNRAVLTCSEQDGSPPEYTWFKDGIVMPTNPKSTRAFSNS  
SYVLNPTTGELVFDPLSASDTGEYSCEARNGYGTPMTSNAVRMEAVERNVGVIVA AVLVTLI  
LLGILVFGIWFAYS RGHFDRTKKGTSSKKVIYSQPSARSEGEFKQTSSFLV

### **Signal sequence:**

amino acids 1-27

### **Transmembrane domain:**

amino acids 238-255

### **N-glycosylation site.**

amino acids 185-189

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 270-274

### **Casein kinase II phosphorylation site.**

amino acids 34-38, 82-86, 100-104, 118-122, 152-156, 154-158,  
193-197, 203-207, 287-291

### **N-myristoylation site.**

amino acids 105-111, 116-122, 158-164, 219-225, 237-243, 256-262

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



3300 Highway 2, Apt 11  
 North York, Ont. M2H 3A7  
 Canada

CAGCGCGTGGCCGGCGCCGCTGTGGGGACAGCATGAGCGGCGGTTGGATGGCGCAGGTTGGA  
GCGTGGCGAACAGGGGCTCTGGGCCTGGCGCTGCTGCTGCTGCTCGGCCTCGGACTAGGCCT  
GGAGGCCGCCGCGAGCCCGCTTTCCACCCCGACCTCTGCCCAGGCCGCAGGCCCCAGCTCAG  
GCTCGTGGCCACCCACCAAGTTCCAGTGCCGCACCAGTGGCTTATGCGTGCCCCCTCACCTGG  
CGCTGCGACAGGGACTTGGACTGCAGCGATGGCAGCGATGAGGAGGAGTGCAGGATTGAGCC  
ATGTACCCAGAAAGGGCAATGCCACCCGCCCCCTGGCCTCCCCTGCCCTGCACCGGCGTCA  
GTGACTGCTCTGGGGGAAGTGAACAAGAACTGCGCAACTGCAGCCGCCTGGCCTGCCTAGCA  
GGCGAGCTCCGTTGCACGCTGAGCGATGACTGCATTCCAATCACGTGGCGCTGCGACGGCCA  
CCCAGACTGTCCCGACTCCAGCGACGAGCTCGGCTGTGGAACCAATGAGATCCTCCCGGAAG  
GGGATGCCACAACCATGGGGCCCCCTGTGACCCTGGAGAGTGTCCCTCTGTGCGGAATGCCACATCCTC  
ACAACCATGGGGCCCCCTGTGACCCTGGAGAGTGTCCCTCTGTGCGGAATGCCACATCCTC  
CTCTGCCGGAGACCAGTCTGGAAGCCCAACTGCCTATGGGGTTATTGCAGCTGCTGCGGTGC  
TCAGTGCAAGCCTGGTCACCGCCACCCTCCTCCTTTTGTCTGGCTCCGAGCCCAGGAGCGC  
CTCCGCCCCTGAGGTTACTGGTGGCCATGAAGGAGTCCCTGCTGCTGTCAGAACAGAAGAC  
CTCGCTGCCCTTGAGGACAAGCACTTGCCACCACCGTCACTCAGCCCTGGGCGTAGCCGGACA  
GGAGGAGAGCAGTGATGCGGATGGGTACCCGGGCACACCAGCCCTCAGAGACCTGAGTTCTT  
CTGGCCACGTGGAACCTCGAACCCGAGCTCCTGCAGAAAGTGGCCCTGGAGATTGAGGGTCCC  
TGGACACTCCCTATGGAGATCCGGGGAGCTAGGATGGGGAACTGCCACAGCCAGAACTGAG  
GGGCTGGCCCCAGGCAGCTCCAGGGGGTAGAACGGCCCTGTGCTTAAGACACTCCCTGCTG  
CCCCGTCTGAGGGTGGCGATTAAAGTTGCTTC

## **FIGURE 46**

MSGGWMAQVGAWRTGALGLALLLLGLGLGLEAAASPLSTPTSAQAAGPSSGSCPPTKFQCR  
TSGLCVPLTWRCRDLDCSDGSDEEEECRIEPC TQKGQCP PPPGLPCPCTGVSDCSGGTDKKL  
RNCSRLACLAGE LRCTLSDDCIPLTWRC DGHPCDSSDELGCGTNEILPEGDATTMGPPVT  
LESVTSLRNATTMGPPVTLESVPSVGNATSSSAGDQSGSPTAYGVIAAAVLSASLVTATLL  
LLSWLRAQERLRPLGLLVAMKESLLLSEQKTSLP

### **Signal sequence:**

amino acids 1-30

### **Transmembrane domain:**

amino acids 230-246

### **N-glycosylation site.**

amino acids 126-130, 195-199, 213-217

### **Casein kinase II phosphorylation site.**

amino acids 84-88, 140-144, 161-165, 218-222

### **N-myristoylation site.**

amino acids 3-9, 10-16, 26-32, 30-36, 112-118, 166-172, 212-218,  
224-230, 230-236, 263-269

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 44-55

### **Leucine zipper pattern.**

amino acids 17-39

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## FIGURE 47

CCCACGCGTCCGGTCTCGCTCGCTCGCGCAGCGGCGGCAGCAGAGGTCGCGCACAGATGCGG  
GTTAGACTGGCGGGGGGAGGAGGCGGAGGAGGGAAGGAAGCTGCATGCATGAGACCCACAGA  
CTCTTGCAAGCTGGATGCCCTCTGTGGATGAAAGATGTATCATGGAATGAACCCGAGCAATG  
GAGATGGATTTCTAGAGCAGCAGCAGCAGCAGCAACCTCAGTCCCCCAGAGACTCTTG  
GCCGTGATCCTGTGGTTTCAGCTGGCGCTGTGCTTCGGCCCTGCACAGCTCACGGGCGGGTT  
CGATGACCTTCAAGTGTGTGCTGACCCCGGCATTCCCGAGAATGGCTTCAGGACCCCGAGCG  
GAGGGGTTTTCTTTGAAGGCTCTGTAGCCCGATTTCACTGCCAAGACGGATTCAAGCTGAAG  
GGCGCTACAAAGAGACTGTGTTTGAAGCATTTTAATGGAACCCTAGGCTGGATCCCAAGTGA  
TAATTCATCTGTGTGCAAGAAGATTGCCGTATCCCTCAAATCGAAGATGCTGAGATTCATA  
ACAAGACATATAGACATGGAGAGAAGCTAATCATCACTTGTCATGAAGGATTCAAGATCCGG  
TACCCCGACCTACACAATATGGTTTTATTATGTGCGCATGATGGAACGTGGAATAATCTGCC  
CATCTGTCAAGGCTGCCTGAGACCTCTAGCCTCTTCTAATGGCTATGTAAACATCTCTGAGC  
TCCAGACCTCCTTCCCGGTGGGGACTGTGATCTCCTATCGCTGCTTTCCCGGATTTAAACTT  
GATGGGTCTGCGTATCTTGAGTGCTTACAAAACCTTATCTGGTCGTCCAGCCCACCCGGTG  
CCTTGCTCTGGAAGCCCAAGTCTGTCCACTACCTCCAATGGTGAGTCACGGAGATTTCTGTCT  
GCCACCCGCGGCCTTGTGAGCGCTACAACCACGGAACCTGTGGTGGAGTTTTACTGCGATCCT  
GGCTACAGCCTCACCAGCGACTACAAGTACATCACCTGCCAGTATGGAGAGTGGTTTTCTTC  
TTATCAAGTCTACTGCATCAAATCAGAGCAAACGTGGCCCAGCACCCATGAGACCCTCCTGA  
CCACGTGGAAGATTGTGGCGTTACGGCAACCAGTGTGCTGCTGGTGCTGCTGCTCGTCATC  
CTGGCCAGGATGTTCCAGACCAAGTTCAAGGCCCACTTTCCCCCAGGGGGCCTCCCCGGAG  
TTCCAGCAGTGACCCTGACTTTGTGGTGGTAGACGGCGTGCCCGTCATGCTCCCGTCCTATG  
ACGAAGCTGTGAGTGGCGGCTTGAGTGCTTAGGCCCGGGTACATGGCCTCTGTGGGCCAG  
GGCTGCCCCTTACCCGTGGACGACCAGAGCCCCCAGCATAACCCGGCTCAGGGGACACGGA  
CACAGGCCCAGGGGAGTCAGAAACCTGTGACAGCGTCTCAGGCTCTTCTGAGCTGCTCCAAA  
GTCTGTATTACCTCCCAGGTGCCAAGAGAGCACCCACCCTGCTTCGGACAACCCTGACATA  
ATTGCCAGCACGGCAGAGGAGGTGGCATCCACCAGCCCAGGCATCCATCATGCCCCACTGGGT  
GTTGTTCTTAAGAAACTGATTGATTAAAAAATTTCCCAAAGTGTCTGAAGTGTCTCTTCAA  
ATACATGTTGATCTGTGGAGTTGATTCTTTCTTCTCTTGGTTTTAGACAAATGTAAACAA  
AGCTCTGATCCTTAAAATTGCTATGCTGATAGAGTGGTGAGGGCTGGAAGCTTGATCAAGTC  
CTGTTTCTTCTTGACACAGACTGATTAAAAATTAAGNAAAAA

## **FIGURE 48**

MYHGMNPSNGDGFLEQQQQQQQPQSPQRLLAVILWFQLALCFGPAQLTGGFDDLQVCADPGI  
PENGFRTPSGGVFFEGSVARFHCQDGFKLKGATKRLCLKHFNGTLGWIPSDNSICVQEDCRI  
PQIEDAEIHNKTYRHGEKLIITCHEGFKIRYPDLHNMVSLCRDDGTWNNLPICQGCLRPLAS  
SNGYVNISELQTSFPVGTVISYRCFPFGFKLDGSAYLECLQNLIWSSSPPRCLALEAQVCPLP  
PMVSHGDFVCHPRPCERYNHGTVVEFYCDPGYSLTSDYKYITCQYGEWFPSYQVYCIKSEQT  
WPSTHETLLTTWKIVAFTATSVLLVLLLVLARMFQTKFKAHFPPRGPPRSSSSDPDFVVVD  
GVPVMLPSYDEAVSGGLSALGPGYMASVGQGCPLPVDDQSPPAYPGSGD TDTGPGESETCDS  
VSGSSELLQSLYSPPRCQESTHPASDNPDIIASTAEVASTSPGIHHAHWVFLRN

### **Signal sequence:**

amino acids 1-41

### **Transmembrane domain:**

amino acids 325-344

### **N-glycosylation site.**

amino acids 104-108, 134-138, 192-196

### **Casein kinase II phosphorylation site.**

amino acids 8-12, 146-150, 252-256, 270-274, 313-317, 362-366,  
364-368, 380-384, 467-471, 468-472

### **N-myristoylation site.**

amino acids 4-10, 61-67, 169-175, 203-209, 387-393, 418-424,  
478-484

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 394-405

Protein Data Bank

## FIGURE 49

CCCACGCGTCCGCTCCGCGCCCTCCCCCGCCTCCCGTGCGGTCCGTCGGTGGCCTAGAGA  
TGCTGCTGCCGCGGTTCAGTTGTCGCGCACGCCTCTGCCCGCCAGCCCGCTCCACCGCCGT  
AGCGCCCGAGTGTCTGGGGGGCGCACCCGAGTCGGGCCATGAGGCCGGGAACCGCGCTACAGG  
CCGTGCTGCTGGCCGTGCTGCTGGTGGGGCTGCGGGCCGCGACGGGTGCGCTGCTGAGTGCC  
TCGGATTGGACCTCAGAGGAGGGCAGCCAGTCTGCCGGGGAGGGACACAGAGGCCTTGTTA  
TAAAGTCATTTACTTCCATGATACTTCTCGAAGACTGAACTTTGAGGAAGCCAAAGAAGCCT  
GCAGGAGGGATGGAGGCCAGCTAGTCAGCATCGAGTCTGAAGATGAACAGAACTGATAGAA  
AAGTTCATTGAAAACCTCTTGCCATCTGATGGTGACTTCTGGATTGGGCTCAGGAGGCGTGA  
GGAGAAACAAAGCAATAGCACAGCCTGCCAGGACCTTTATGCTTGACTGATGGCAGCATAT  
CACAATTTAGGAACTGGTATGTGGATGAGCCGTCCTGCGGCAGCGAGGTCTGCGTGGTCATG  
TACCATCAGCCATCGGCACCCGCTGGCATCGGAGGCCCTACATGTTCCAGTGGAATGATGA  
CCGGTGCAACATGAAGAACAATTTCAATTTGCAAATATTCTGATGAGAAACCAGCAGTTCCTT  
CTAGAGAAGCTGAAGGTGAGGAAACAGAGCTGACAACACCTGTACTTCCAGAAGAAACACAG  
GAAGAAGATGCCAAAAAACATTTAAAGAAAGTAGAGAAGCTGCCTTGAATCTGGCCTACAT  
CCTAATCCCCAGCATTCCCCCTTCTCCTCCTCCTTGTGGTCACCACAGTTGTATGTTGGGTTT  
GGATCTGTAGAAAAAGAAAACGGGAGCAGCCAGACCCTAGCACAAAGAAGCAACACACCATC  
TGGCCCTCTCCTCACCAGGGAAACAGCCCGGACCTAGAGGTCTACAATGTCATAAGAAAACA  
AAGCGAAGCTGACTTAGCTGAGACCCGGCCAGACCTGAAGAATATTTCAATTCCGAGTGTGTT  
CGGGAGAAGCCACTCCCGATGACATGTCTTGTGACTATGACAACATGGCTGTGAACCCATCA  
GAAAGTGGGTTTGTGACTCTGGTGAGCGTGGAGAGTGGATTTGTGACCAATGACATTTATGA  
GTTCTCCCCAGACCAAATGGGGAGGAGTAAGGAGTCTGGATGGGTGGAAAATGAAATATATG  
GTTATTAGGACATATAAAAACTGAACTGACAACAATGGAAAAGAAATGATAAGCAAAATC  
CTCTTATTTTCTATAAGGAAAATACACAGAAGGTCTATGAACAAGCTTAGATCAGGTCCTGT  
GGATGAGCATGTGGTCCCCACGACCTCCTGTTGGACCCCCACGTTTTGGCTGTATCCTTTAT  
CCCAGCCAGTCATCCAGCTCGACCTTATGAGAAGGTACCTTGCCCAGGTCTGGCACATAGTA  
GAGTCTCAATAAATGTCACCTGGTTGGTTGTATCTAACTTTTAAGGGACAGAGCTTTACCTG  
GCAGTGATAAAGATGGGCTGTGGAGCTTGAAAACCACCTCTGTTTTCTTTGCTCTATACAG  
CAGCACATATTATCATACAGACAGAAAATCCAGAATCTTTTCAAAGCCCACATATGGTAGCACAG  
GTTGGCCTGTGCATCGGCAATTCTCATATCTGTTTTTTTCAAAGAATAAAATCAAATAAAGA  
GCAGGAAAAAAAAA

## **FIGURE 50**

MRPGTALQAVLLAVLLVGLRAATGRLLSASDLRLGGQPVCRGGTQRPCYKVIYFHDTSRRL  
NFEEAKEACRRDGGQLVSI ESEDEQKLI EKFIENLLPSDGD FWIGLRRREEKQSNSTACQDL  
YAWTDGSISQFRNWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKY  
SDEKPAVPSREAEGEETELTTPVLPEETQEEDAKKTFKESREAALNLAYILIPSIPLLLLVLV  
VTTVVCWVWICRKRKREQPD PSTKKQHTIWPSPHQGN SPDLEVYNVIRKQSEADLAETRPDL  
KNISFRVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGRSKES  
GWVENEIYGY

### **Signal sequence:**

amino acids 1-21

### **Transmembrane domain:**

amino acids 235-254

### **N-glycosylation site.**

amino acids 117-121, 312-316

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 296-300

### **Casein kinase II phosphorylation site.**

amino acids 28-32, 30-34, 83-87, 100-104, 214-218, 222-226,  
299-303, 306-310, 323-327

### **N-myristoylation site.**

amino acids 18-24, 37-43, 76-82, 146-152

## FIGURE 51

GGGGTCTCCCTCAGGGCCGGGAGGCACAGCGGTCCCTGCTTGCTGAAGGGCTGGATGTACGC  
ATCCGCAGGTTCCCGCGGACTTGGGGGCGCCCGCTGAGCCCCGGCGCCCGCAGAAGACTTGT  
GTTTGCCTCCTGCAGCCTCAACCCGGAGGGCAGCGAGGGCCTACCACCATGATCACTGGTGT  
G TTCAGCATGCGCTTGTGGACCCAGTGGGCGTCCTGACCTCGCTGGCGTACTGCCTGCACC  
AGCGGCGGGTGGCCCTGGCCGAGCTGCAGGAGGCCGATGGCCAGTGTCCGGTTCGACCGCAGC  
CTGCTGAAGTTGAAAATGGTGCAGGTCGTGTTTCGACACGGGGCTCGGAGTCCTCTCAAGCC  
GCTCCCGCTGGAGGAGCAGGTAGAGTGGAACCCCCAGCTATTAGAGGTCCCACCCCAAATC  
AGTTTGATTACACAGTCACCAATCTAGCTGGTGGTCCGAAACCATATTCTCCTTACGACTCT  
CAATACCATGAGACCACCCTGAAGGGGGGCATGTTTGCTGGGCAGCTGACCAAGGTGGGCAT  
GCAGCAAATGTTTGCCTTGGGAGAGAGACTGAGGAAGAACTATGTGGAAGACATTCCCTTTC  
TTTCACCAACCTTCAACCCACAGGAGGTCTTTATTTCGTTCCACTAACATTTTTTCGGAATCTG  
GAGTCCACCCGTTGTTTGCTGGCTGGGCTTTTTCCAGTGTCAGAAAGAAGGACCCATCATCAT  
CCACACTGATGAAGCAGATT CAGAAGTCTTGATCCCAACTACCAAAGCTGCTGGAGCCTGA  
GGCAGAGAACCAGAGGCCGGAGGCAGACTGCCTCTTTACAGCCAGGAATCTCAGAGGATTTG  
AAAAAGGTGAAGGACAGGATGGGCATTGACAGTAGTGATAAAGTGGACTTCTTCATCCTCCT  
GGACAACGTGGCTGCCGAGCAGGCACACAACCTCCCAAGCTGCCCCATGCTGAAGAGATTTG  
CACGGATGATCGAACAGAGAGCTGTGGACACATCCTTGTACATACTGCCCAAGGAAGACAGG  
GAAAGTCTTCAGATGGCAGTAGGCCCATTCCTCCACATCCTAGAGAGCAACCTGCTGAAAGC  
CATGGACTCTGCCACTGCCCCCGACAAGATCAGAAAGCTGTATCTCTATGCGGCTCATGATG  
TGACCTTCATACCGCTCTTAATGACCCTGGGGATTTTTTGACCACAAATGGCCACCGTTTGCT  
GTTGACCTGACCATGGAACCTTACCAGCACCTGGAATCTAAGGAGTGGTTTGTGCAGCTCTA  
TTACCACGGGAAGGAGCAGGTGCCGAGAGGTTGCCCTGATGGGCTCTGCCCCGCTGGACATGT  
TCTTGAATGCCATGTCAGTTTATACCTTAAGCCCAGAAAAATACCATGCACTCTGCTCTCAA  
ACTCAGGTGATGGAAGTTGGAATGAAGAGTAACTGATTTATAAAAGCAGGATGTGTTGATT  
TTAAAATAAAGTGCCTTTATACAATG

## **FIGURE 52**

MITGVFSMRLWTPVGVLTSLAYCLHQRRVALAELQEADGQCPVDRSLLKLKMOVVFRHGAR  
SPLKPLPLEEQVEWNPQLLEVPPQTQFDYTVTNLAGGPKPYSPYDSQYHETTLKGGMFAGQL  
TKVGMQQMFALGERLRKNYVEDIPFLSPTFNPQEVFIRSTNIFRNLESTRCLLAGLFQCQKE  
GPIIIHTDEADSEVLYPNYQSCWSLRQRTGRRRQTASLQPGISEDLKKVKDRMGIDSSDKVD  
FFILLDNVAAEQAHNLPSCPMLKRFARMIEQRAVDTSLYILPKEDRESLQMAVGPFHLHILES  
NLLKAMDSATAPDKIRKLYLYAAHDVTFIPLLMTLGIFDHKWPPFAVDLTMELYQHLESKEW  
FVQLYYHGKEQVPRGCPDGLCPDLMFLNAMS VYTLSP EKYHALCSQTQVMEVGNEE

### **Signal sequence:**

amino acids 1-23

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 218-222

### **Casein kinase II phosphorylation site.**

amino acids 87-91, 104-108, 320-324

### **Tyrosine kinase phosphorylation site.**

amino acids 280-288

### **N-myristoylation site.**

amino acids 15-21, 117-123, 118-124, 179-185, 240-246, 387-393

### **Amidation site.**

amino acids 216-220

### **Leucine zipper pattern.**

amino acids 10-32

### **Histidine acid phosphatases phosphohistidine signature.**

amino acids 50-65



## FIGURE 53

CTCCTCTTAACATACTTGCAGCTAAAACTAAATATTGCTGCTTGGGGACCTCCTTCTAGCCT  
TAAATTTTACAGCTCATCACCTTCACCTGCCTTGGTCAATGGCTCTGCTATTCTCCTTGATCCTT  
GCCATTTTGCACCAGACCTGGATTCTTAGCGTCTCCATCTGGAGTGC GGCTGGTGGGGGGCCT  
CCACCGCTGTGAAGGGCGGGTGGAGGTGGAACAGAAAGGCCAGTGGGGCACCGTGTGTGATG  
ACGGCTGGGACATTAAGGACGTGGCTGTGTTGTGCCGGGAGCTGGGCTGTGGAGCTGCCAGC  
GGAACCCCTAGTGGTATTTTGTATGAGCCACCAGCAGAAAAAGAGCAAAAGGTCCTCATCCA  
ATCAGTCAGTTGCACAGGAACAGAAGATACATTGGCTCAGTGTGAGCAAGAAGAAGTTTATG  
ATTGTTACATGATGAAGATGCTGGGGCATCGTGTGAGAACCCAGAGAGCTCTTTCTCCCCA  
GTCCCAGAGGGTGTGAGGCTGGCTGACGGCCCTGGGCATTGCAAGGGACGCGTGGAAGTGAA  
GCACCAGAACCAGTGGTATACCGTGTGCCAGACAGGCTGGAGCCTCCGGGGCCGCAAAGGTGG  
TGTGCCGGCAGCTGGGATGTGGGAGGGCTGTACTGACTCAAAAACGCTGCAACAAGCATGCC  
TATGGCCGAAAACCCATCTGGCTGAGCCAGATGTCATGCTCAGGACGAGAAGCAACCCCTTCA  
GGATTGCCCTTCTGGGCCTTGGGGGAAGAACACCTGCAACCATGATGAAGACACGTGGGTCG  
AATGTGAAGATCCCTTTGACTTGAGACTAGTAGGAGGAGACAACCTCTGCTCTGGGCGACTG  
GAGGTGCTGCACAAGGGCGTATGGGGCTCTGTCTGTGATGACAACCTGGGGAGAAAAGGAGGA  
CCAGGTGGTATGCAAGCAACTGGGCTGTGGGAAGTCCCTCTCTCCCTCCTTCAGAGACCGGA  
AATGCTATGGCCCTGGGGTTGGCCGCATCTGGCTGGATAATGTTGCTCAGGGGAGGAG  
CAGTCCCTGGAGCAGTGCCAGCACAGATTTTGGGGGTTTCACGACTGCACCCACCAGGAAGA  
TGTGGCTGTCTCTGCTCAGTGTAGGTGGGCATCATCTAATCTGTTGAGTGCCTGAATAGAA  
GAAAAACACAGAAGAAGGGAGCATTTACTGTCTACATGACTGCATGGGATGAACACTGATCT  
TCTTCTGCCCTTGGACTGGGACTTATACTTGGTGCCCTGATTCTCAGGCCCTTCAGAGTTGG  
ATCAGAACTTACAACATCAGGTCTAGTTCTCAGGCCATCAGACATAGTTTGGAACTACATCA  
CCACCTTTCCTATGTCTCCACATTGCACACAGCAGATTCCCAGCCTCCATAATTGTGTGTAT  
CAACTACTTAAATACATTCTCACACACACACACACACACACACACACACACACACATA  
CACCATTTGTCCTGTTTTCTCTGAAGAACTCTGACAAAATACAGATTTTGGTACTGAAAGAGA  
TTCTAGAGGAACGGAATTTTAAGGATAAATTTTCTGAATTGGTTATGGGGTTTCTGAAATTG  
GCTCTATAATCTAATTAGATATAAAATTCTGGTAACTTTATTTACAATAATAAAGATAGCAC  
TATGTGTTCAAA

## FIGURE 54

MALLFSLILAICTRPGFLASPSGVRLVGGLHRCEGRVEVEQKGQWGTVCDDGWDIKDVAVLC  
RELGCGAASGTPSGILYEPPEAEKEQKVLIQSVSCTGTEDTLAQCEQEVEYDCSHDEDAGASC  
ENPESSFSPVPEGVRLADGPGHCKGRVEVKHQNQWYTVCTGWSLRAAKVVCRLGCGRAVL  
TQKRCNKHAYGRKPIWLSQMSCSGREATLQDCPSGPWGKNTCNHDEDTWVECEDPFDLRLVG  
GDNLCSGRLEVLHKGWVGSVCDDNWGEKEDQVVCKQLGCGKSLSPSFRDRKCYGPGVGRIWL  
DNVRCSGEEQSLEQCQHRFWGFHDCTHQEDVAVICSV

### **Signal sequence:**

amino acids 1-15

### **Casein kinase II phosphorylation site.**

amino acids 47-51, 97-101, 115-119, 209-213, 214-218, 234-238,  
267-271, 294-298, 316-320, 336-340

### **N-myristoylation site.**

amino acids 29-35, 43-49, 66-72, 68-74, 72-78, 98-104, 137-143,  
180-186, 263-269, 286-292

### **Amidation site.**

amino acids 196-200

### **Speract receptor repeated domain signature.**

amino acids 29-67, 249-287

Sequence "CH3360"

## FIGURE 55

ACTGCACTCGGTTCTATCGATTGAATTCCTCGGGGATCCTCTAGAGATCCCTCGACCTCGAC  
CCACGCGTCCGCGGACGCGTGGGCGGACGCGTGGGCCGGCTACCAGGAAGAGTCTGCCGAAG  
GTGAAGGCCATGGACTTCATCACCTCCACAGCCATCCTGCCCTGCTGTTCCGGCTGCCTGGG  
CGTCTTCGGCCTCTTCCGGCTGCTGCAGTGGGTGCGCGGGAAGGCCTACCTGCGGAATGCTG  
TGGTGGTGATCACAGGCGCCACCTCAGGGCTGGGCAAAGAATGTGCAAAAGTCTTCTATGCT  
GCGGGTGCTAAACTGGTGCTCTGTGGCCGGAATGGTGGGGCCCTAGAAGAGCTCATCAGAGA  
ACTTACCGCTTCTCATGCCACCAAGGTGCAGACACACAAGCCTTACTTGGTGACCTTCGACC  
TCACAGACTCTGGGGCCATAGTTGCAGCAGCAGCTGAGATCCTGCAGTGCTTTGGCTATGTC  
GACATACTTGTCACAATGCTGGGATCAGCTACCGTGGTACCATCATGGACACCACAGTGGA  
TGTGGACAAGAGGGTCATGGAGACAACTACTTTGGCCCAGTTGCTCTAACGAAAGCACTCC  
TGCCCTCCATGATCAAGAGGAGGCAAGGCCACATTGTCGCCATCAGCAGCATCCAGGGCAAG  
ATGAGCATTCCTTTTCGATCAGCATATGCAGCCTCCAAGCACGCAACCCAGGCTTTCTTTGA  
CTGTCTGCGTGCCGAGATGGAACAGTATGAAATTGAGGTGACCGTCATCAGCCCCGGCTACA  
TCCACACCAACCTCTCTGTAAATGCCATCACCGCGGATGGATCTAGGTATGGAGTTATGGAC  
ACCACCACAGCCCAGGGCCGAAGCCCTGTGGAGGTGGCCCAGGATGTTCTTGCTGCTGTGGG  
GAAGAAGAAGAAAGATGTGATCCTGGCTGACTTACTGCCTTCCTTGGCTGTTTATCTTCGAA  
CTCTGGCTCCTGGGCTCTTCTTCAGCCTCATGGCCTCCAGGGCCAGAAAAGAGCGGAAATCC  
AAGAACTCCTAGTACTCTGACCAGCCAGGGCCAGGGCAGAGAAGCAGCACTCTTAGGCTTGC  
TTACTCTACAAGGGACAGTTGCATTTGTTGAGACTTTAATGGAGATTTGTCTCACAAGTGGG  
AAAGACTGAAGAAACACATCTCGTGCAGATCTGCTGGCAGAGGACAATCAAAAACGACAACA  
AGCTTCTTCCAGGGTGAGGGGAAACACTTAAGGAATAAATATGGAGCTGGGGTTTAACACT  
AAAAACTAGAAATAAACATCTCAACAGTAAAAAAAAAAAAAAAAAGGGCGGCCGCGACTCTAG  
AGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCACTTGTTTATTGCAGCTTATAATGGTTAC

## **FIGURE 56**

MDFITSTAILPLLFGCLGVFGLFRLQLQWVRGKAYLRNAVVVITGATSGLGKECAKVFYAAGA  
KLVLCGRNGGALEELIRELTASHATKVQTHKPYLVTFDLTDSGAIVAAAAEILQCFGYVDIL  
VNNAGISYRGTIMDTTVDDVKRVMETNYFGPVALTKALLPSMIKRRQGHIVAIISSIQGKMSI  
PFRSAYAASKHATQAFFDCLRAEMEQYEIEVTVISPGYIHTNLSVNAITADGSRYGVMDTTT  
AQGRSPVEVAQDVLAAVGKKKKDVILADLLPSLAVYLRTLAPGLFFSLMASRARKERKSKNS

### **Signal sequence:**

amino acids 1-21

### **Transmembrane domain:**

amino acids 104-120, 278-292

### **N-glycosylation site.**

amino acids 228-232

### **Glycosaminoglycan attachment site.**

amino acids 47-51

### **Casein kinase II phosphorylation site.**

amino acids 135-139, 139-143, 253-257

### **Tyrosine kinase phosphorylation site.**

amino acids 145-153, 146-153

### **N-myristoylation site.**

amino acids 44-50, 105-111, 238-244, 242-248, 291-297

### **Amidation site.**

amino acids 265-269

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 6-17

bioRxiv preprint doi: <https://doi.org/10.1101/2020.03.10.333333>; this version posted March 10, 2020. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## FIGURE 57

CCCACGCGTCCGCTGGTGTAGATCGAGCAACCCTCTAAAAGCAGTTTAGAGTGGTAAAAAA  
AAAAAAAAACACACCAAACGCTCGCAGCCACAAAAGGGATGAAATTTCTTCTGGACATCCTC  
CTGCTTCTCCCGTTACTGATCGTCTGCTCCCTAGAGTCCTTCGTGAAGCTTTTTATTCTCTAA  
GAGGAGAAAATCAGTCACCGGCGAAATCGTGCTGATTACAGGAGCTGGGCATGGAATTGGGA  
GACTGACTGCCTATGAATTTGCTAAACTTAAAAGCAAGCTGGTTCTCTGGGATATAAATAAG  
CATGGACTGGAGGAAACAGCTGCCAAATGCAAGGGACTGGGTGCCAAGGTTTCATACCTTTGT  
GGTAGACTGCAGCAACCGAGAAGATATTTACAGCTCTGCAAAGAAGGTGAAGGCAGAAATTG  
GAGATGTTAGTATTTTAGTAAATAATGCTGGTGTAGTCTATACATCAGATTTGTTTGCTACA  
CAAGATCCTCAGATTGAAAAGACTTTTGAAGTTAATGTACTTGCACATTTCTGGACTACAAA  
GGCATTTCCTTCTGCAATGACGAAGAATAACCATGGCCATATTGTCACTGTGGCTTCGGCAG  
CTGGACATGTCTCGGTCCCCCTTCTTACTGGCTTACTGTTCAAGCAAGTTTGCTGCTGTTGGA  
TTTCATAAACTTTGACAGATGAACTGGCTGCCTTACAAATAACTGGAGTCAAAACAACATG  
TCTGTGTCTTAATTTCTGTAAACACTGGCTTCATCAAAAATCCAAGTACAAGTTTGGGACCCA  
CTCTGGAACCTGAGGAAGTGGTAAACAGGCTGATGCATGGGATTCTGACTGAGCAGAAGATG  
ATTTTTATTCCATCTTCTATAGCTTTTTTAACAACATTGGAAAGGATCCTTCCTGAGCGTTT  
CCTGGCAGTTTTTAAACGAAAAATCAGTGTTAAGTTTGATGCAGTTATTGGATATAAAATGA  
AAGCGCAATAAGCACCTAGTTTTCTGAAAAGTATTACCAGGTTTAGGTTGATGTCATCTA  
ATAGTGCCAGAATTTTAATGTTTGAAGTTCTGTTTTTTCTAATTATCCCCATTTCTTCAATA  
TCATTTTTGAGGCTTTGGCAGTCTTCATTTACTACCACTTGTTCTTTAGCCAAAAGCTGATT  
ACATATGATATAAACAGAGAAATACCTTTAGAGGTGACTTTAAGGAAAATGAAGAAAAAGAA  
CCAAAATGACTTTATTAAAATAATTTCCAAGATTATTTGTGGCTCACCTGAAGGCTTTGCAA  
AATTTGTACCATAACCGTTTATTTAACATATATTTTTATTTTTTGATTGCACTTAAATTTTGT  
ATAATTTGTGTTTCTTTTTCTGTTCTACATAAAATCAGAACTTCAAGCTCTCTAAATAAAA  
TGAAGGACTATATCTAGTGGTATTTTACAATGAATATCATGAACTCTCAATGGGTAGGTTTC  
ATCCTACCCATTGCCACTCTGTTTCTGAGAGATACCTCACATTCCAATGCCAAACATTTCT  
GCACAGGGAAGCTAGAGGTGGATACACGTGTTGCAAGTATAAAAGCATCACTGGGATTTAAG  
GAGAATTGAGAGAATGTACCCACAAATGGCAGCAATAATAAATGGATCACACTTAAAAAAA  
AA  
AAA

2065-11-03-01

amino acids 1-19

amino acids 30-34, 283-287

amino acids 52-56, 95-99, 198-202, 267-271

amino acids 43-49, 72-78, 122-128, 210-216

## FIGURE 59

CCCACGCGTCCGCGGACGCGTGCGTGGGTCGACTAGTTCTAGATCGCGAGCGGCCGCCCGCGGGCTC  
AGGGAGGAGCACCGACTGCGCCGCACCCCTGAGAGATGGTTGGTGCCATGTGGAAGGTGATTG  
TTTCGCTGGTCTGTGATGCCTGGCCCCCTGTGATGGGCTGTTTCGCTCCCTATACAGAAGT  
GTTTCCATGCCACCTAAGGGAGACTCAGGACAGCCATTATTTCTCACCCCTTACATTGAAGC  
TGGGAAGATCCAAAAAGGAAGAGAATTGAGTTTGGTTCGGCCCTTTCCAGGACTGAACATGA  
AGAGTTATGCCGGCTTCTCACCGTGAATAAGACTTACAACAGCAACCTCTTCTTCTGGTTC  
TTCCCAGCTCAGATACAGCCAGAAGATGCCCCAGTAGTTCTCTGGCTACAGGGTGGGCCGGG  
AGGTTTCATCCATGTTTGGACTCTTTGTGGAACATGGGCCTTATGTTGTACAAGTAACATGA  
CCTTGCGTGACAGAGACTTCCCCTGGACCACAACGCTCTCCATGCTTTACATTGACAATCCA  
GTGGGCACAGGCTTCAGTTTTACTGATGATACCCACGGATATGCAGTCAATGAGGACGATGT  
AGCACGGGATTTATACAGTGCCTAATTCAGTTTTTCCAGATATTTCTGAATATAAAAAATA  
ATGACTTTTTATGTCACTGGGGAGTCTTATGCAGGGAAATATGTGCCAGCCATTGCACACCTC  
ATCCATTCCCTCAACCCTGTGAGAGAGGTGAAGATCAACCTGAACGGAATTGCTATTGGAGA  
TGGATATTCTGATCCCGAATCAATTATAGGGGGCTATGCAGAATTCCTGTACCAAATTGGCT  
TGTTGGATGAGAAGCAAAAAAAGTACTTCCAGAAGCAGTGCCATGAATGCATAGAACACATC  
AGGAAGCAGAAGTGGTTTGAGGCCTTTGAAATACTGGATAAACTACTAGATGGCGACTTAAC  
AAGTGATCCTTCTTACTTCCAGAATGTTACAGGATGTAGTAATTAATACTTTTGTGGGT  
GCACGGAACCTGAGGATCAGCTTTACTATGTGAAATTTTTGTCACTCCAGAGGTGAGACAA  
GCCATCCACGTGGGGAATCAGACTTTTAATGATGGAACCTATAGTTGAAAAGTACTTGCGAGA  
AGATACAGTACAGTCAGTTAAGCCATGGTTAACTGAAATCATGAATAATTATAAGGTTCTGA  
TCTACAATGGCCAACTGGACATCATCGTGGCAGCTGCCCTGACAGAGCGCTCCTTGATGGGC  
ATGGACTGGAAAGGATCCCAGGAATACAAGAAGGCAGAAAAAAAGTTTGGAAGATCTTTAA  
ATCTGACAGTGAAGTGGCTGGTTACATCCGGCAAGCGGGTGAATTCATCAGGTAATTATTC  
GAGGTGGAGGACATATTTTACCCTATGACCAGCCTCTGAGAGCTTTTGACATGATTAATCGA  
TTCATTTATGGAAGGATGGGATCCTTATGTTGGATAAACTACCTTCCCAAAGAGAACAT  
CAGAGGTTTTTCATTGCTGAAAAGAAAATCGTAAAAACAGAAAATGTCATAGGAATAAAAAA  
TTATCTTTTCATATCTGCAAGATTTTTTTTCATCAATAAAAAATTATCCTTGAAACAAGTGAGC  
TTTTGTTTTTGGGGGAGATGTTTACTACAAAATTAACATGAGTACATGAGTAAGAATTACA  
TTATTTAACTTAAAGGATGAAAGGTATGGATGATGTGACACTGAGACAAGATGTATAAATGA  
AATTTTAGGGTCTTGAATAGGAAGTTTAAATTTCTTCTAAGAGTAAGTGAAGTGCAGTTG  
TAACAAACAAAGCTGTAACATCTTTTTCTGCCAATAACAGAAGTTTGGCATGCCGTGAAGGT  
GTTTGGAATATTATTGGATAAGAATAGCTCAATTATCCCAAATAAATGGATGAAGCTATAA  
TAGTTTTGGGGAAAAGATTCTCAAATGTATAAAGTCTTAGAACAAAAGAATTCTTTGAAATA  
AAAATATTATATATAAAAGTAAAAA

## **FIGURE 60**

MVGAMWKVIVSLVLLMPGPCDGLFRSLYRSVSMPPKGD SGQPLFLTPYIEAGKIQKGRELSL  
VGPFPG LNMKSYAGFLTVNKTYNSNLFFWFFPAQIQPEDAPVVLWLQGGPGGSSMFGLFVEH  
GPYVVTSNMTLRDRDFPWTTTTLSMLYIDNPVGTGFSFTDDTHGYAVNEDDVARDLYSALI QF  
FQIFPEYKNND FYVTGESYAGKYVPAIAHLIHSLNPVREVKINLNGIAIGDGYSDPESIIGG  
YAEFLYQIGLLDEKQKKYFQKQCHECIEHIRKQNWFEAFEILDKLLDGDLTSDPSYFQNVTG  
CSNYYNFLRCTEPEDQLYYVKFLSLPEVRQAIHVGNQTFNDGTIVEKYLREDTVQSVKPWLT  
EIMNNYKVLIIYNGQLDIIVAAALTERSLMGMDWKGSQEYKKA EKKVWKIFKSDSEVAGYIRQ  
AGDFHQVIIRGGGHILPYDQPLRAFDMINRFIYGKGWDPYVG

### **Signal sequence:**

amino acids 1-22

### **N-glycosylation site.**

amino acids 81-85, 132-136, 307-311, 346-350

### **Casein kinase II phosphorylation site.**

amino acids 134-138, 160-164, 240-244, 321-325, 334-338, 348-352,  
353-357, 424-428

### **Tyrosine kinase phosphorylation site.**

amino acids 423-432

### **N-myristoylation site.**

amino acids 22-28, 110-116, 156-162, 232-238

### **Serine carboxypeptidases, serine active site.**

amino acids 200-208

### **Crystallins beta and gamma 'Greek key' motif signature.**

amino acids 375-391

FIGURE 60



## FIGURE 61

CGAGGGCTTTTCCGGCTCCGGAATGGCACATGTGGGAATCCCAGTCTTGTGGCTACAACAT  
TTTTCCCTTTCCTAACAGTTCTAACAGCTGTTCTAACAGCTAGTGATCAGGGGTTCTTCTT  
GCTGGAGAAGAAAGGGCTGAGGGCAGAGCAGGGCACTCTCACTCAGGGTGACCAGCTCCTTG  
CCTCTCTGTGGATAACAGAGCATGAGAAAGTGAAGAGATGCAGCGGAGTGAGGTGATGGAAG  
TCTAAAATAGGAAGGAATTTTGTGTGCAATATCAGACTCTGGGAGCAGTTGACCTGGAGAGC  
CTGGGGGAGGGCCTGCCTAACAGCTTTCAAAAAACAGGAGCGACTTCCACTGGGCTGGGAT  
AAGACGTGCCGGTAGGATAGGGAAGACTGGGTTTAGTCCTAATATCAAATTGACTGGCTGGG  
TGAACCTTCAACAGCCTTTTAACCTCTCTGGGAGATGAAAACGATGGCTTAAGGGGGCCAGAAA  
TAGAGATGCTTTGTAAAATAAAATTTTAAAAAAGCAAGTATTTTATAGCATAAAGGCTAGA  
GACCAAAATAGATAACAGGATTCCCTGAACATTCCCTAAGAGGGAGAAAGTATGTTAAAAATA  
GAAAAACCAAAATGCAGAAGGAGGAGACTCACAGAGCTAAACCAGGATGGGGACCCTGGGTC  
AGGCCAGCCTCTTTGCTCCTCCCGGAAATTATTTTGGTCTGACCACTCTGCCTTGTGTTTT  
GCAGAATCATGTGAGGGCCAACCGGGGAAGGTGGAGCAGATGAGCACACACAGGAGCCGTCT  
CCTCACCGCCGCCCTCTCAGCATGGAACAGAGGCAGCCCTGGCCCCGGGCCCTGGAGGTGG  
ACAGCCGCTCTGTGGTCTGCTCTCAGTGGTCTGGGTGCTGCTGGCCCCCCCCAGCAGCCGGC  
ATGCCTCAGTTCAGCACCTTCCACTCTGAGAATCGTGACTGGACCTTCAACCACTTGACCGT  
CCACCAAGGGACGGGGGCCGTCTATGTGGGGGCCATCAACCGGGTCTATAAGCTGACAGGCA  
ACCTGACCATCCAGGTGGCTCATAAGACAGGGGCCAGAAGAGGACAACAAGTCTCGTTACCCG  
CCCCTCATCGTGCAGCCCTGCAGCGAAGTGCTCACCTCACCAACAATGTCAACAAGCTGCT  
CATCATTGACTACTCTGAGAACCGCCTGCTGGCCTGTGGGAGCCTCTACCAGGGGTCTGCA  
AGCTGCTGCGGTGGATGACCTCTTCATCCTGGTGGAGCCATCCACAAGAAGGAGCACTAC  
CTGTCCAGTGTCACAAGACGGGCACCATGTACGGGGTGATTGTGCGCTCTGAGGGTGAGGA  
TGGCAAGCTCTTCATCGGCACGGCTGTGGATGGGAAGCAGGATTACTTCCCGACCCTGTCCA  
GCCGGAAGCTGCCCCGAGACCCTGAGTCTCAGCCATGCTCGACTATGAGCTACACAGCGAT  
TTTGTCTCCTCTCTCATCAAGATCCCTTCAGACACCCTGGCCCTGGTCTCCCACTTTGACAT  
CTTCTACATCTACGGCTTTGCTAGTGGGGGCTTTGTCTACTTTCTCACTGTCCAGCCCCGAGA  
CCCCTGAGGGTGTGGCCATCAACTCCGCTGGAGACCTCTTCTACACCTCACGCATCGTGCGG  
CTCTGCAAGGATGACCCCAAGTTCCACTCATACTGTCCCTGCCCTTCGGCTGCACCCGGGC  
CGGGGTGGAATACCGCCTCCTGCAGGCTGCTTACCTGGCCAAGCCTGGGGACTCACTGGCCC  
AGGCCTTCAATATCACAGCCAGGACGATGTACTCTTTGCCATCTTCTCCAAAGGGCAGAAG  
CAGTATCACCAACCCGCCGATGACTCTGCCCTGTGTGCCCTTCCCTATCCGGGCCATCAACTT  
GCAGATCAAGGAGCGCCTGCAGTCTGTACCAGGGCGAGGGCAACCTGGAGCTCAACTGGC  
TGCTGGGGAAGGACGTCCAGTGACGAAGGCGCCTGTCCCCATCGATGATAACTTCTGTGGA  
CTGGACATCAACCAGCCCCTGGGAGGCTCAACTCCAGTGGAGGGCCTGACCCTGTACACCAC  
CAGCAGGGACCGCATGACCTCTGTGGCCTCCTACGTTTACAACGGCTACAGCGTGGTTTTTG  
TGGGGACTAAGAGTGGCAAGCTGAAAAAGGTAAGAGTCTATGAGTTCAGATGCTCCAATGCC  
ATTACCTCCTCAGCAAAGAGTCCCTCTTGGAAGGTAGCTATTGGTGGAGATTTAACTATAG  
GCAACTTTATTTTCTTGGGGAACAAAGGTGAATGGGGAGGTAAGAAGGGGTTAATTTTGTG  
ACTTAGCTTCTAGCTACTTCCCTCCAGCCATCAGTCATTGGGTATGTAAGGAATGCAAGCGTA  
TTTCAATATTTCCCAAACCTTTAAGAAAAAAGCTTTAAGAAGGTACATCTGCAAAAGCAA

## FIGURE 62

MGTLGQASLFAPPGNYFWS DHSALCF AESCEGQPGKVEQMSTHR S RLLTAAPLSMEQRQPWP  
RALEVDSRSVVL LSVVWVLLAPPAAGMPQFSTFHSEN RDWTFNHLTVHQGTGAVYVGAINRV  
YKLTGNLTIQVAHKTGPEEDNKSRY PPLIVQPCSEVLTLTNNVNKLLIIDYSENRL LACGSL  
YQGVCKLLRLDDLFI LVEPSHKKEHYLSSVNKTGTMYGVIVRSEGEDGKLFIGTAVDGKQDY  
FPTLSSRKLPRDP ESSAMLDYELHSD FVSSLIKIPSDTLALVSHFDIFYIYGFASGGFVYFL  
TVQPETPEGVAINSAGDLFYTSRIVRLCKDDPKFHSYVSLPFGCTRAGVEYRLLQAAYLAKP  
GDSLAAQAFNITSQDDVLF AIFSKGQKQYHHPDD SALCAFP IRAINLQIKERLQSCYQGEGN  
LELNWLLGKDVQCTKAPVPIDDNFCGLDINQPLGGSTPVEGLTLYTTSRDRMTSVASYVYNG  
YSVVFVGTKSGKLKKVRVYEFRC SNAIHLLSKESLLEGSYWWRFNYRQLYFLGEQR

### **Signal sequence:**

amino acids 1-32

### **Transmembrane domain:**

amino acids 71-87

### **N-glycosylation site.**

amino acids 130-134, 145-149, 217-221, 381-385

### **Casein kinase II phosphorylation site.**

amino acids 139-143, 229-233, 240-244, 291-295, 324-328, 383-387,  
384-388, 471-475, 481-485, 530-534

### **N-myristoylation site.**

amino acids 220-226, 319-325, 353-359, 460-466, 503-509

## FIGURE 63

AGGCTCCCGCGCGCGGCTGAGTGCGGACTGGAGTGGGAACCCGGGTCCCGCGCTTAGAGAACACGCGATGACCA  
CGTGGAGCCTCCGGCGGAGGCCGCGCCGCACGCTGGGACTCCTGCTGCTGGTCTTGGGCTTCTGGTGTCTCC  
GCAGGCTGGACTGGAGCACCTTGGTCCCTCTGCGGCTCCGCCATCGACAGCTGGGGCTGCAGGCCAAGGGCTGGA  
ACTTCATGCTGGAGGATTCCACCTTCTGGATCTTCGGGGGCTCCATCCACTATTTCCGTGTGCCAGGGAGTACT  
GGAGGGACCGCTGCTGAAGATGAAGCCTGTGGCTTGAACACCCTCACCACCTATGTTCCGTGGAACCTGCATG  
AGCCAGAAAGAGGCAAATTTGACTTCTCTGGGAACCTGGACCTGGAGGCCTTCGTCTGTATGGCCGAGAGATCG  
GGCTGTGGGTGATTCTGCGTCCAGGCCCCCTACATCTGCAGTGAGATGGACCTCGGGGGCTTGCCAGCTGGCTAC  
TCCAAGACCCTGGCATGAGGCTGAGGACAACCTTACAAGGGCTTCAACGAAGCAGTGGACCTTTATTTTGACCACC  
TGATGTCCAGGGTGGTGCCACTCCAGTACAAGCGTGGGGGACCTATCATTGCCGTGCAGGTGGAGAATGAATATG  
GTTCTTATAATAAAGACCCCGCATACATGCCCTACGTCAAGAAGGCACTGGAGGACCGTGGCATTGTGGAACCTGC  
TCCTGACTTCAGACAACAAGGATGGGCTGAGCAAGGGGATTTGTCCAGGGAGTCTTGGCCACCATCAACTTGCAGT  
CAACACACGAGCTGCAGTACTGACCACCTTTCTCTTCAACGTCCAGGGGACTCAGCCCAAGTGGTATGGAGT  
ACTGGACGGGGTGGTTTGAATCGTGGGGAGGCCCTCACAATATCTTGGATTCTTCTGAGGTTTTGAAAACCGTGT  
CTGCCATTGTGGACGCCGGCTCCTCCATCAACCTCTACATGTTCCACGGAGGCACCAACTTTGGCTTCATGAATG  
GAGCCATGCACCTCCATGACTACAAGTCAGATGTCACCAGCTATGACTATGATGCTGTGCTGACAGAAGCCGGCG  
ATTACACGGCCAAGTACATGAAGCTTCGAGACTTCTTCGGCTCCATCTCAGGCATCCCTCTCCCTCCCCACCTG  
ACCTTCTTCCCAAGATGCCGTATGAGCCCTTAACGCCAGTCTTGTACCTGTCTCTGTGGGACGCCCTCAAGTACC  
TGGGGGAGCCAATCAAGTCTGAAAAGCCCATCAACATGGAGAACCTGCCAGTCAATGGGGGAAATGGACAGTCTCT  
TCGGGTACATTCTCTATGAGACCAGCATCACCTCGTCTGGCATCTCAGTGGCCACGTGCATGATCGGGGGCAGG  
TGTTTGTGAACACAGTATCCATAGGATTCTTGGACTACAAGACAACGAAGATTGCTGTCCCCCTGATCCAGGGTT  
ACACCGTGCTGAGGATCTTGGTGGAGAATCGTGGGCGAGTCAACTATGGGGAGAATATTGATGACCAGCGCAAAG  
GCTTAATTGGAAATCTCTATCTGAATGATTCAACCTGAAAACTTCAGAATCTATAGCCTGGATATGAAGAAGA  
GCTTCTTTCAGAGGTTCCGGCTGGACAAATGGNGTTCCCTCCCAGAAACACCCACATTACCTGCTTTCTTCTTGG  
GTAGCTTGTCCATCAGCTCCACGCCTTGTGACACCTTTCTGAAGCTGGAGGGCTGGGAGAAGGGGGTTGTATTCA  
TCAATGGCCAGAACCTTGGACGTTACTGGAACATTGGACCCAGAAGACGCTTTACCTCCCAGGTCCCTGGTTGA  
GCAGCGGAATCAACCAGGTATCGTTTTTGGAGAGACGATGGCGGGCCCTGCATTACAGTTCACGGAAACCCCCC  
ACCTGGGCAGGAACCAAGTACATTAAGTGAGCGGTGGCACCCCTCCTGCTGGTGCCAGTGGGAGACTGCCGCCTC  
CTCTTGACCTGAAGCCTGGTGGCTGCTGCCCCACCCCTCACTGCAAAGCATCTCCTTAAGTAGCAACCTCAGGG  
ACTGGGGGCTACAGTCTGCCCCGTGTCTCAGCTCAAAACCTAAGCCTGCAGGGAAGGTGGGATGGCTCTGGGCC  
TGGCTTTGTGATGATGGCTTTCTACAGCCCTGCTCTTGTGCCGAGGCTGTGCGGCTGTCTTAGGGTGGGAGC  
AGCTAATCAGATCGCCAGCCTTTGGCCCTCAGAAAAAGTGCTGAAACGTGCCCTTGACCCGGACGTCACAGCCC  
TGCGAGCATCTGCTGGACTCAGGCGTGCTCTTTGCTGGTTCCTGGGAGGCTTGGCCACATCCCTCATGGCCCCAT  
TTTATCCCCGAAATCCTGGGTGTGTACCAAGTGTAGAGGGTGGGGAAGGGGTGTCTACCTGAGCTGACTTTGTT  
CTTCTTCAACCTTCTGAGCCTTCTTTGGGATTCTGGAAGGAACCTCGGCGTGAGAAACATGTGACTTCCCCTT  
TCCCTTCCCACTCGTGTCTTCCACAGGGTGACAGGCTGGGCTGGAGAAACAGAAATCCTCACCTGCGTCTTCC  
CAAGTTAGCAGGTGTCTCTGGTGTTCAGTGAGGAGGACATGTGAGTCTTGGCAGAAGCCATGGCCCATGTCTGCA  
CATCCAGGGAGGAGGACAGAAGGCCAGCTCACATGTGAGTCTTGGCAGAAGCCATGGCCCATGTCTGCACATCC  
AGGGAGGAGGACAGAAGGCCAGCTCACATGTGAGTCTTGGCAGAAGCCATGGCCCATGTCTGCACATCCAGGGA  
GGAGGACAGAAGGCCAGCTCACATGTGAGTCTTGGCAGAAGCCATGGCCCATGTCTGCACATCCAGGGAGGAGG  
ACAGAAGGCCAGCTCAGTGGCCCCCGCTCCCCACCCCCACGCCGAACAGCAGGGGACAGAGCAGCCCTCCTTC  
GAAGTGTGTCCAAGTCCGATTTGAGCCTTGTTCTGGGGCCAGCCCAACACCTGGCTTGGGCTCACTGTCTGA  
GTTGCAGTAAAGCTATAACCTTGAATCAAA

## FIGURE 64

MTTWSLRRRPPARTLGLLLLLVVLGFLVLRRLDWSTLVPLRLRHRQLGLQAKGWNFMLEDSTFW  
IFGGSIHVFRVPREYWRDRLLKMKACGLNTLTITYVPWNLHEPERGKFDFSGNLDLEAFVLMA  
AEIGLWVILRPGPYICSEMDLGGLPSWLLQDPGMRLRTTYKGFTEAVDLYFDHLMSRVVPLQ  
YKRGGPPIIAVQVENEYGSYNKDPAYMPYVKKALEDRGIVELLTSDNKDGLSKGIVQGVLAT  
INLQSTHELQLLTTFLFNVQGTQPKMVMYWTGWFDWSGPPHNILDSSEVLKTVSAIVDAGS  
SINLYMFHGGTNFGFMNGAMHFHDYKSDVTSYDYDAVLTEAGDYTAKYMKLRDFFGSISGIP  
LPPPPDLLPKMPYEPLTPVLYLSLWDALKYLGEPIKSEKPINMENLPVNGGNGQSFGYILYE  
TSITSSGILSGHVHVRGQVFVNTVSIQFLDYKTTKIAVPLIQGYTVLRILVENRGRVNYGEN  
IDDQRKGLIGNLYLNDSPKKNFRIYSLDMKKSFFQRFGLDKWXSLETPPTLPAPFLGSLSLIS  
STPCDTFLKLEGWEKGVVFINGQNLGRYWNIGPQKTLYLPGPWLSGGINQVIVFEETMAGPA  
LQFTETPHLGRNQYIK

### **Signal sequence:**

amino acids 1-27

### **Casein kinase II phosphorylation site.**

amino acids 141-148, 253-257, 340-344, 395-399, 540-544, 560-564

### **N-myristoylation site.**

amino acids 146-152, 236-242, 240-246, 244-250, 287-293, 309-315,  
320-326, 366-372, 423-429, 425-431, 441-447, 503-509, 580-586

## FIGURE 65

GGGGACGCGGAGCTGAGAGGCTCCGGGCTAGCTAGGTGTAGGGGTGGACGGGTCCCAGGACC  
CTGGTGAGGGTTCTCTACTTGGCCTTCGGTGGGGGTCAAGACGCAGGCACCTACGCCAAAGG  
GGAGCAAAGCCGGGCTCGGCCCCGAGGCCCCCAGGACCTCCATCTCCCAATGTTGGAGGAATC  
CGACACGTGACGGTCTGTCCGCCGTCTCAGACTAGAGGAGCGCTGTAAACGCCCATGGCTCCC  
AAGAAGCTGTCTTGCCTTCGTTCCCTGCTGCTGCCGCTCAGCCTGACGCTACTGCTGCCCCA  
GGCAGACACTCGGTTCGTTCTAGTGGATAGGGGTTCATGACCGGTTTCTCCTAGACGGGGCCC  
CGTTCCGCTATGTGTCTGGCAGCCTGCACTACTTTCGGGTACCGCGGGTGCTTTGGGCGGAC  
CGGCTTTTGAAGATGCGATGGAGCGGCCTCAACGCCATACAGTTTTATGTGCCCTGGAACTA  
CCACGAGCCACAGCCTGGGGTCTATAACTTTAATGGCAGCCGGGACCTCATTGCCTTTCTGA  
ATGAGGCAGCTCTAGCGAACCTGTTGGTCATACTGAGACCAGGACCTTACATCTGTGCAGAG  
TGGGAGATGGGGGGTCTCCCATCCTGGTTGCTTCGAAAACCTGAAATTCATCTAAGAACCCTC  
AGATCCAGACTTCCTTGCCGCAGTGGACTCCTGGTTCAAGGTCTTGCTGCCCAAGATATATC  
CATGGCTTTATCACAATGGGGGCAACATCATTAGCATTGAGGTGGAGAATGAATATGGTAGC  
TACAGAGCCTGTGACTTCAGCTACATGAGGCACCTGGCTGGGCTCTTCCGTGCACTGCTAGG  
AGAAAAGATCTTGCTCTTCACCACAGATGGGCCTGAAGGACTCAAGTGTGGCTCCCTCCGGG  
GACTCTATACCACTGTAGATTTTGGCCCAGCTGACAACATGACCAAAATCTTTACCTTGCTT  
CGGAAGTATGAACCCCATGGGCCATTGGTAAACTCTGAGTACTACACAGGCTGGCTGGATTA  
CTGGGGCCAGAATCACTCCACACGGTCTGTGTGTCAGCTGTAACCAAAGGACTAGAGAACATGC  
TCAAGTTGGGAGCCAGTGTGAACATGTACATGTTCCATGGAGGTACCAACTTTGGATATTGG  
AATGGTGCCGATAAGAAGGGACGCTTCCTTCCGATTACTACCAGCTATGACTATGATGCACC  
TATATCTGAAGCAGGGGACCCACACCTAAGCTTTTTGCTCTTCGAGATGTCATCAGCAAGT  
TCCAGGAAGTTCTTTGGGACCTTTACCTCCCCCGAGCCCCAAGATGATGCTTGGACCTGTG  
ACTCTGCACCTGGTTGGGCATTTACTGGCTTTCTAGACTTGCTTTGCCCCCGTGGGCCCCAT  
TCATTCAATCTTGCCAATGACCTTTGAGGCTGTCAAGCAGGACCATGGCTTCATGTTGTACC  
GAACCTATATGACCCATAACATTTTGGAGCCAACACCATTCTGGGTGCCAAATAATGGAGTC  
CATGACCGTGCTTATGTGATGGTGGATGGGGTGTTCCAGGGTGTTGTGGAGCGAAATATGAG  
AGACAAACTATTTTTGACGGGGAAACTGGGGTCCAAACTGGATATCTTGGTGGAGAACATGG  
GGAGGCTCAGCTTTGGGTCTAACAGCAGTGACTTCAAGGGCCTGTTGAAGCCACCAATTCTG  
GGCAAACAATCCTTACCCAGTGGATGATGTTCCCTCTGAAAATTGATAACCTTGTGAAGTG  
GTGGTTTCCCCTCCAGTTGCCAAAATGGCCATATCCTCAAGCTCCTTCTGGCCCCACATTCT  
ACTCCAAAACATTTCCAATTTTAGGCTCAGTTGGGGACACATTTCTATATCTACCTGGATGG  
ACCAAGGGCCAAGTCTGGATCAATGGGTTTAACTTGGGCGGTAAGTGGACAAAGCAGGGGCC  
ACAACAGACCCTCTACGTGCCAAGATTCTGCTGTTTCTAGGGGAGCCCTCAACAAAATTA  
CATTGCTGGAACTAGAAGATGTACCTCTCCAGCCCCAAGTCCAATTTTTGGATAAGCCTATC  
CTCAATAGCACTAGTACTTTGCACAGGACACATATCAATTCCCTTTCAGCTGATACTAGG  
TGCCTCTGAACCAATGGAGTTAAGTGGGCACCTGAAAGGTAGGCCGGGCATGGTGGCTCATGC  
CTGTAATCCCAGCACTTTGGGAGGCTGAGACGGGTGGATTACCTGAGGTGAGGACTTCAAGA  
CCAGCCTGGCCAACATGGTGAACCCCGTCTCCACTAAAAATACAAAAATTAGCCGGGCGTG  
ATGGTGGGCACCTCTAATCCCAGCTACTTGGGAGGCTGAGGGCAGGAGAATTGCTTGAATCC  
AGGAGGCAGAGGTTGCAGTGAAGTGGAGGTTGTACCACTGCACTCCAGCCTGGCTGACAGTGA  
GACACTCCATCTCAAAAAAAAAAAAA

## **FIGURE 66**

MAPKKLSCLRSLLLPLSLTLLLPQADTRSFVVDGRHDFLLDGAPFRYVSGSLHYFRVPRVL  
WADRLLKMRWSGLNAIQFYVPWNYHEPQPGVYNFNGSRDLIAFLNEAALANLLVILRPGPYI  
CAEWEMGGLPSWLLRKPEIHLRTSDPDFLAAVDSWFKVLLPKIYPWLYHNGGNIISIQVENE  
YGSYRACDFS YMRHLAGLFRALLGEKILLFTTDGPEGLKCGSLRGLYTTVDFGPADNMTKIF  
TL LRKYEPHGPLVNSEYYTGWLDYWGQNHSTRSVSAVTKGLENMLKLGASVNM MYFHGGTNF  
GYWNGADKKGRFLPITTSYDYDAPISEAGDPTPKLFALRDVISKFQEVPLGPLPPSPKMML  
GPVTLHLVGHLLAFLDLLCPRGPIHSILPMTFEAVKQDHGFMLYRTYMTHTIFEPTPFWVPN  
NGVHDRAYVMVDGVFQGVVERNMRDKLFLTGKLGSKLDILVENMGRLSFGSNSSDFKGLLKP  
PILGQTILTQWMMFPLKIDNLVKWWFPLQLPKWPYPQAPSGPTFYSKTFPILGSVGD TFLYL  
PGWTKGQVWINGFNLGRYWTKQGPQQTLYVPRFLLFPRGALNKITLLELEDVPLQPQVQFLD  
KPILNSTSTLHRTHINSLSADTLSASEPMELSGH

### **Signal sequence:**

amino acids 1-27

### **N-glycosylation site.**

amino acids 97-101, 243-247, 276-280, 486-490, 625-629

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 4-8

### **Casein kinase II phosphorylation site.**

amino acids 148-152, 234-238, 327-331, 423-427, 469-473, 550-554,  
603-607, 644-648

### **Tyrosine kinase phosphorylation site.**

amino acids 191-198

### **N-myristoylation site.**

amino acids 131-137, 176-182, 188-194, 203-209, 223-229, 227-233,  
231-237, 274-280, 296-300, 307-313, 447-453, 484-490

## FIGURE 67

GCTTTGAACACGTCTGCAAGCCCAAAGTTGAGCATCTGATTGGTTATGAGGTATTTGAGTGC  
ACCCACAATATGGCCTTACATGTTGAAAAAGCTTCTCATCAGTTACATATCCATTATTTGTGT  
TTATGGCTTTATCTGCCTCTACACTCTCTTCTGGTTATTCAGGATACCTTTGAAGGAATATT  
CTTTCGAAAAAGTCAGAGAAGAGAGCAGTTTTAGTGACATTCCAGATGTCAAAAACGATTTT  
GCGTTCCTTCTTCACATGGTAGACCAGTATGACCAGCTATATTCCAAGCGTTTTGGTGTGTT  
CTTGTCAGAAGTTAGTGAAAATAAACTTAGGGAAATTAGTTTGAACCATGAGTGGACATTTG  
AAAAACTCAGGCAGCACATTTACGCAACGCCAGGACAAGCAGGAGTTGCATCTGTTTCATG  
CTGTGCGGGGTGCCCGATGCTGTCTTTGACCTCACAGACCTGGATGTGCTAAAGCTTGAAC  
AATTCCAGAAGCTAAAATTCCTGCTAAGATTTCTCAAATGACTAACCTCCAAGAGCTCCACC  
TCTGCCACTGCCCTGCAAAAGTTGAACAGACTGCTTTTTAGCTTTCTTCGCGATCACTTGAGA  
TGCCTTCACGTGAAGTTCACTGATGTGGCTGAAATTCCTGCCTGGGTGTATTTGCTCAAAAA  
CCTTCGAGAGTTGTACTTAATAGGCAATTTGAACTCTGAAAACAATAAGATGATAGGACTTG  
AATCTCTCCGAGAGTTGCGGCACCTTAAGATTCTCCACGTGAAGAGCAATTTGACCAAAGTT  
CCCTCCAACATTACAGATGTGGCTCCACATCTTACAAAGTTAGTCATTATAATGACGGCAC  
TAAACTCTTGGTACTGAACAGCCTTAAGAAAATGATGAATGTCGCTGAGCTGGAACCTCAGA  
ACTGTGAGCTAGAGAGAATCCCACATGCTATTTTCAGCCTCTCTAATTTACAGGAACTGGAT  
TTAAAGTCCAATAACATTCGCACAATTGAGGAAATCATCAGTTTCCAGCATTTAAAACGACT  
GACTTGTTTTAAAATTATGGCATAACAAAATTGTTACTATTCTCCCTCTATTACCCATGTCA  
AAAACCTTGGAGTCACTTTATTTCTCTAACAACAAGCTCGAATCCTTACCAGTGGCAGTATTT  
AGTTTACAGAACTCAGATGCTTAGATGTGAGCTACAACAACATTTCAATGATTCCAATAGA  
AATAGGATTGCTTCAGAACCTGCAGCATTTGCATATCACTGGGAACAAAGTGGACATTCTGC  
CAAAACAATTGTTTTAAATGCATAAAGTTGAGGACTTTGAATCTGGGACAGAACTGCATCACC  
TCACTCCCAGAGAAAGTTGGTCAGCTCTCCAGCTCACTCAGCTGGAGCTGAAGGGGAACTG  
CTTGGAACCGCCTGCCAGCCCAGCTGGGCCAGTGTGCGATGCTCAAGAAAAGCGGGCTTGTTG  
TGGAAGATCACCTTTTTTGATACCCTGCCACTCGAAGTCAAAGAGGCATTGAATCAAGACATA  
AATATTCCCTTTGCAAATGGGATTTTAAACTAAGATAATATATGCACAGTGATGTGCAGGAAC  
AACTTCCTAGATTGCAAGTGCTCACGTACAAGTTATTACAAGATAATGCATTTTAGGAGTAG  
ATACATCTTTTAAAATAAAACAGAGAGGATGCATAGAAGGCTGATAGAAGACATAACTGAAT  
GTTCAATGTTTGTAGGGTTTTAAGTCATTCATTTCCAAATCATTTTTTTTTTTCTTTTGGGG  
AAAGGGAAGGAAAAATTATAATCACTAATCTTGGTTCTTTTTTAAATTGTTTGTAACCTGGAT  
GCTGCCGCTACTGAATGTTTACAAATTGCTTGCCTGCTAAAGTAAATGATTAAATTGACATT  
TTCTTACTAAAAAAAAAAAAAAAAA

## **FIGURE 68**

MAYMLKKLLISYISIIICVYGFICLYTLFWLFRIPLKEYSFEKVREESSFSDIPDVKNDF AFL  
LHMVDQYDQLYSKRFGVFLSEVSENKLR EISLNHEWTFEKL RQHISR NAQDKQELHLFMLS G  
VPDAVFDLTDL DVLKLELIPEAKIPAKISQMTNLQELHLCHCPAKVEQTAFSFLRDHLRCLH  
VKFTDVAEIPAWVYLLKNLRELYLIGNLNSENNKMIGLES LREL RHLKILHVKS NLTKVPSN  
ITDVAPHLTKLVIHNDGTKLLVLNSLKKMMNVAE LELQNC ELERIPHAIFSLSNLQELDLKS  
NNIRTIEEIIISFQHLKRLTCLKLWHNKIVTIPPSITHVKNLES LYFSNNKLES LPVAVFSLQ  
KL RCLDVS YNNISMIP IEIGLLQNLQHLHITGNKV DILPKQLFKCIKLRTLNLGQNCITSLP  
EKVGQLSQLTQLELKGNC LDRLPAQLGQCRMLKKSGLVVEDHLFDTLPLEVKEALNQDINIP  
FANGI

### **Signal sequence:**

amino acids 1-20

### **N-glycosylation site.**

amino acids 241-245, 248-252, 383-387

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 326-330

### **Casein kinase II phosphorylation site.**

amino acids 48-52, 133-137, 226-230, 315-319, 432-436, 444-448

### **Tyrosine kinase phosphorylation site.**

amino acids 349-355, 375-381

### **N-myristoylation site.**

amino acids 78-84, 124-130, 212-218, 392-398



## FIGURE 69

CCCACGCGTCCGGCCCTTCTCTCTGGACTTTGCATTTCCATTCCCTTTTCATTGACAACTGACTTTTTTTTATTTCT  
TTTTTTCCATCTCTGGGCCAGCTTGGGATCCTAGGCCGCCCTGGGAAGACATTTGTGTTTTACACACATAAGGAT  
CTGTGTTTTGGGGTTTCTTCTTCTCCCTGACATTGGCATTGCTTAGTGTTGTGTGGGGAGGGAGACCACGTGG  
GCTCAGTGCTTGCTTGCATTATCTGCCTAGGTACATCGAAGTCTTTTGACCTCCATACAGTGATTATGCCTGTC  
ATCGCTGGTGGTATCCTGGCGGCCTTGCTCCTGCTGATAGTTGTCTGTCTGTCTTTACTTCAAATACACAAC  
GCGCTAAAAGCTGCAAAGGAACCTGAAGCTGTGGCTGTAAAAAATCACAACCAGACAAGGTGTGGTGGGCCAAG  
AACAGCCAGGCCAAAACCATTGCCACGGAGTCTTGTCTGCCCCGAGTGCTGTGAAGGATATAGAATGTGTGCC  
AGTTTTGATTCCCTGCCACCTTGCTGTTGCGACATAAATGAGGGCCTCTGAGTTAGGAAAGGCTCCCTTCTCAA  
GCAGAGCCCTGAAGACTTCAATGATGTCAATGAGGCCACCTGTTTGTGATGTGCAGGCACAGAAGAAAGGCACAG  
CTCCCCATCAGTTTCATGGAAAATAACTCAGTGCCTGCTGGGAACCCAGCTGCTGGAGATCCCTACAGAGAGCTTC  
CACTGGGGGCAACCCTTCCAGGAAGGAGTTGGGGAGAGAGAACCCTCACTGTGGGGAATGCTGATAAACCACTCA  
CACAGCTGCTCTATTCTCACACAAATCTACCCCTTGCGTGGCTGGAACCTGACGTTTCCCTGGAGGTGTCCAGAAA  
GCTGATGTAACACAGAGCCTATAAAAGCTGTGCGTCTTAAGGCTGCCAGCGCTTGCCAAAATGGAGCTTGTA  
AGAAGGCTCATGCCATTGACCTCTTAATTCCTCTCCTGTTTGGCGGAGCTGACAATGGCGGAGGCTGAAGGCAAT  
GCAAGCTGCACAGTCAGTCTAGGGGGTGCCAATATGGCAGAGACCCACAAAGCCATGATCCTGCAACTCAATCCC  
AGTGAGAACTGCACCTGGACAATAGAAAGACCAGAAAACAAAAGCATCAGAATTATCTTTTCTATGTCCAGCTT  
GATCCAGATGGAAGCTGTGAAAAGTGAAAACATTAAAGTCTTTGACGGAACCTCCAGCAATGGGCCTCTGCTAGGG  
CAAGTCTGCAGTAAAAACGACTATGTTCTGTATTTGAATCATCATCCAGTACATTGACGTTTCAAATAGTTACT  
GACTCAGCAAGAATTCAAAGAACTGTCTTTGTCTTCTACTACTTCTTCTCTCTAACATCTCTATTCCAACTGT  
GGCGGTTTACCTGGATACCTTGGAAAGGATCCTTACCAGCCCCAATTACCCAAAGCCGCATCCTGAGCTGGCTTAT  
TGTGTGTGGCACATACAAGTGGAGAAAGATTACAAGATAAACTAAACTTCAAAGAGATTTTCTTAGAAATAGAC  
AAACAGTGCAAATTTGATTTTCTTGCCATCTATGATGGCCCCCTCCACCAACTCTGGCCTGATTGGACAAGTCTGT  
GGCCGTGTGACTCCACCTTCCAATCGTCATCAAACCTCTCTGACTGTCTGTGTTGTCTACAGATTATGCCAATTCT  
TACCGGGGATTTTCTGCTTCTACACCTCAATTTATGCAGAAAACATCAACACTACATCTTTAACTTGCTCTTCT  
GACAGGATGAGAGTTATTATAAGCAAATCCTACCTAGAGGCTTTTAACTCTAATGGGAATAACTTGCAACTAAAA  
GACCCAACTTGACAGACCAAAATTATCAAATGTTGTGGAATTTTCTGTCCCTCTTAATGGATGTGGTACAATCAGA  
AAGGTAGAAGATCAGTCAATTACTTACACCAATATAATCACCTTTTCTGCATCCTCAACTTCTGAAGTGATCACC  
CGTCAGAAACAACTCCAGATTATTGTGAAGTGTGAAATGGGACATAATTCTACAGTGAGATAATATACATAACA  
GAAGATGATGTAATACAAAGTCAAAATGCACTGGGCAAAATATAACACCAGCATGGCTCTTTTGAATCCAATTCA  
TTTGAAAAGACTATACTTGAATCACCATATTATGTGGATTTGAACCAAACCTCTTTTGTTCAGTTAGTCTGCAC  
ACCTCAGATCCAAATTTGGTGGTGTCTTGTGATACCTGTAGAGCCTCTCCACCTCTGACTTTGCATCTCCAACC  
TACGACCTAATCAAGAGTGGATGTAGTCGAGATGAACTTGTAAAGGTGTATCCCTTATTTGGACACTATGGGAGA  
TTCCAGTTTAAATGCCTTTAAATTCCTTGAGAAGTATGAGCTCTGTGTATCTGCAGTGTAAGTTTTGATATGTGAT  
AGCAGTGACCACCAGTCTCGCTGCAATCAAGTTGTGTCTCCAGAAGCAAACGAGACATTTCTCATATAAATGG  
AAAACAGATTCCATCATAGGACCCATTCTGTCTGAAAAGGGATCGAAGTGCAAGTGGCAATTCAAGGATTTTCA  
GAAACACATGCGGAAGAACTCCAAACCAGCCTTTCAACAGTGTGCATCTGTTTTCTTTCATGGTTCTAGCTCTG  
AATGTGGTGAAGTGTAGCGACAATCACAGTGAGGCATTTTGTAAATCAACGGGCAGACTACAAATACCAGAAGCTG  
CAGAACTATTAACTAACAGGTCCAACCTAAGTGAGACATGTTTCTCCAGGATGCCAAAGGAAATGTACCTCGT  
GGCTACACATATTATGAATAAATGAGGAAGGGCTGAAAGTGACACACAGGCCTGCATGTAAAAAAA

## **FIGURE 70**

MELVRRRLMPLTLLILSCLAELTMAEAEAGNASCTVSLGGANMAETHKAMILQLNPSENCTWTI  
ERPENKSIRIIIFSIVQLDPDGSCSENIKVFDGTSSNGPLLQVCSKNDYVPVFESSSSTLT  
FQIVTDSARIQRTVFVFFYYFFSPNISIPNCGGYLDTLEGSFTSPNYPKPHPELAYCVWHIQV  
EKDYKIKLNFKEIFLEIDKQCKFDLAIYDGPSTNSGLIGQVCGRVTPTFESSSNSLTVVLS  
TDYANSYRGFSASYTSIYAENINTTSLTCSSDRMRVIIISKSYLEAFNSNGNNLQLKDPTCRP  
KLSNVVEFSVPLNGCGTIRKVEDQSITYTNIITFSASSTSEVITRQKQLQIIVKCEMGHNST  
VEIIYITEDDVIQSQNALGKYNTSMALFESNSFEKTILESPYYVDLNQTLFVQVSLHTSDPN  
LVVFLDTCRASPTSDFASPTYDLIKSGCSRDETCKVYPLFGHYGRFQFNAFKFLRSMSSVYL  
QCKVLICDSSDHQSRCNQGCVSRSKRDISSYKWKTDIIIGPIRLKRDRSASGNSGFQHETHA  
EETPNQPFNSVHLFSFMVLALNVVTVATITVRHFVNQRADYKYQKLQNY

### **Signal sequence:**

amino acids 1-24

### **Transmembrane domain:**

amino acids 571-586

### **N-glycosylation site.**

amino acids 29-33, 57-61, 67-71, 148-152, 271-275, 370-374,  
394-398, 419-423

### **Casein kinase II phosphorylation site.**

amino acids 22-26, 108-112, 289-293, 348-352, 371-375, 379-383,  
408-412, 463-467, 520-524, 556-560

### **Tyrosine kinase phosphorylation site.**

amino acids 172-180, 407-415, 407-416, 519-528

### **N-myristoylation site.**

amino acids 28-34, 38-44, 83-89, 95-101, 104-110, 226-232

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 7-18

[illegible]

GACGGGAAGGAACAGACGCTCCCGAGGCGGGGAGCTGCAGAGAGGACAGCCGCCCTGCGCCG  
GGACATGCGGCCCCAGGAGCTCCCCAGGCTCGCGTTCCCGTTGCTGCTGTTGCTGTTGCTGCTG  
TGCTGCCGCCGCCGCGCTGCCCTGCCACAGCGCCACGCGCTTCGACCCACCTGGGAGTCC  
CTGGACGCCCCGCCAGCTGCCCGCGTGGTTTTGACCAGGCCAAGTTCGGCATCTTCATCCACTG  
GGGAGTGTTTTTCCGTGCCCAGCTTCGGTAGCGAGTGGTTCTGGTGGTATTGGCAAAGGAAA  
AGATACCGAAGTATGTGGAATTTATGAAAGATAATTACCTCCTAGTTTCAAATATGAAGAT  
TTTGGACCACTATTTACAGCAAAATTTTTTAATGCCAACCAAGTGGGCAGATATTTTTTCAGG  
CTCTGGTGCCAAATACATTGTCTTAACTTCCAAACATCATGAAGGCTTTACCTTGTGGGGGT  
CAGAATATTCGTGGAACCTGGAATGCCATAGATGAGGGGCCAAGAGGGACATTGTCAAGGAA  
CTTGAGGTAGCCATTAGGAACAGAAGTACCTGCGTTTTGGACTGTACTATTCCCTTTTTTGA  
ATGGTTTCATCCGCTCTTCCCTTGAGGATGAATCCAGTTCATTCCATAAGCGGCAATTTCCAG  
TTTCTAAGACATTGCCAGAGCTCTATGAGTTAGTGAACAACTATCAGCCTGAGGTTCTGTGG  
TCGGATGGTGACGGAGGAGCACCGGATCAATACTGGAACAGCACAGGCTTCTTGGCCTGGTT  
ATATAATGAAAGCCAGTTCGGGGCACAGTAGTCACCAATGATCGTTGGGGAGCTGGTAGCA  
TCTGTAAGCATGGTGGCTTCTATACCTGCAGTGATCGTTATAACCCAGGACATCTTTTGCCA  
CATAAATGGGAAAACCTGCATGACAATAGACAAACTGTCCTGGGGCTATAGGAGGGAAGCTGG  
AATCTCTGACTATCTTACAATTGAAGAATTGGTGAAGCAACTTGTAGAGACAGTTTCATGTG  
GAGGAAATCTTTTGATGAATATTGGGCCCCACACTAGATGGCACCATTCTGTAGTTTTTTGAG  
GAGCGACTGAGGCAAGTGGGGTCTTGGCTAAAAGTCAATGGAGAAGCTATTTATGAAACCTA  
TACCTGGCGATCCAGAATGACACTGTCACCCCAGATGTGTGGTACACATCCAAGCCTAAAG  
AAAAATTAGTCTATGCCATTTTTCTTAAATGGCCACATCAGGACAGCTGTTCCCTTGGCCAT  
CCCAAAGCTATTCTGGGGGCAACAGAGGTGAAACTACTGGGCCATGGACAGCCACTTAAGT  
GATTTCTTTGGAGCAAAATGGCATTATGGTAGAACTGCCACAGCTAACCATTTCATCAGATGC  
CGTGTAATGGGGCTGGGCTCTAGCCCTAACTAATGTGATCTTAAAGTGCAGCAGAGTGGCTG  
ATGCTGCAAGTTATGTCTAAGGCTAGGAACTATCAGGTGTCTATAATTGTAGCACATGGAGA  
AAGCAATGTAAACTGGATAAGAAAATTATTTGGCAGTTCAGCCCTTTCCCTTTTTTCCCCTA  
AATTTTTCTTAAATTACCCATGTAACCATTTTAACTCTCCAGTGCACTTTGCCATTAAAGTC  
TCTTCACATTGATTTGTTTCCATGTGTGACTCAGAGGTGAGAATTTTTTTCACATTATAGTAG  
CAAGGAATTGGTGGTATTATGGACCGAACTGAAAATTTTATGTTGAAGCCATATCCCCCATG  
ATTATATAGTTATGCATCACTTAATATGGGGATATTTTCTGGGAAATGCATTGCTAGTCAAT  
TTTTTTTTTGTGCCAACATCATAGAGTGTATTTACAAAATCCTAGATGGCATAGCCTACTACA  
CACCTAATGTGTATGGTATAGACTGTTGCTCCTAGGCTACAGACATATACAGCATGTTACTG  
AATACTGTAGGCAATAGTAACAGTGGTATTTGTATATCGAAACATATGGAAACATAGAGAAG  
GTACAGTAAAAATACTGTAAAATAAATGGTGCACCTGTATAGGGCACTTACCACGAATGGAG  
CTTACAGGACTGGAAGTTGCTCTGGGTGAGTCAGTGAGTGAATGTGAAGGCCTAGGACATTA  
TTGAACACTGCCAGACGTTATAAATACTGTATGCTTAGGCTACACTACATTTATAAAAAAAA  
GTTTTTCTTTCTTCAATTATAAATTAACATAAGTGTACTGTAACTTTACAAACGTTTTTAATT  
TTTAAACCTTTTTGGCTCTTTTGTAATAACACTTAGCTTAAACATAAACTCATTTGTGCAA  
ATGTAA

## **FIGURE 72**

MRPQELPRLAFPLLLLLLLLLLPPPPCPAHSATRFDPTWESLDARQLPAWFDQAKFGIFIHWG  
VFSVPSFGSEWFWWYQKEKIPKYVEFMKDNYPSPFKYEDFGPLFTAKFFNANQWADIFQAS  
GAKYIVLTSKHHEGFTLWGSEYSWNWNAIDEGPKRDIVKELEVAIRNRTDLRFGLYYSLEW  
FHPLFLEDESSSFHKRQFPVSKTLPELYELVNNYQPEVLWSDGDGGAPDQYWNSTGFLAWLY  
NESPVRGTVVTNDRWGAGSICKHGGFYTCSDRYNPGHLLPHKWENCMTIDKLSWGYRREAGI  
SDYLTIEELVKQLVETVSCGGNLLMNIGPTLDGTISVVFEERLRQVGSWLKVNGEAIYETYT  
WRSQNDTVTPDVWYTSKPKEKLVYAIFLKWPTSGQLFLGHPKAILGATEVKLLGHGQPLNWI  
SLEQNGIMVELPQLTIHQMPCKWGWALALTNVI

### **Signal sequence:**

amino acids 1-28

### **N-glycosylation site.**

amino acids 171-175, 239-243, 377-381

### **Casein kinase II phosphorylation site.**

amino acids 32-36, 182-186, 209-213, 227-231, 276-280, 315-319,  
375-375

### **Tyrosine kinase phosphorylation site.**

amino acids 361-369, 389-397

### **N-myristoylation site.**

amino acids 143-149, 178-184, 255-261, 272-278, 428-434

### **Leucine zipper pattern.**

amino acids 410-432

### **Alpha-L-fucosidase putative active site.**

amino acids 283-295

protein sequence

## FIGURE 73

AGCAGGGAAATCCGGATGTCTCGGTTATGAAGTGGAGCAGTGAGTGTGAGCCTCAACATAGT  
TCCAGAACTCTCCATCCGGACTAGTTATTGAGCATCTGCCTCTCATATCACCAGTGGCCATC  
TGAGGTGTTTCCCTGGCTCTGAAGGGGTAGGCACGATGGCCAGGTGCTTCAGCCTGGTGTG  
CTTCTCACTTCCATCTGGACCACGAGGCTCCTGGTCCAAGGCTCTTTGCGTGCAGAAGAGCT  
TTCCATCCAGGTGTCATGCAGAATTATGGGGATCACCCCTTGTGAGCAAAAAGGCGAACCAGC  
AGCTGAATTTACACAGAAGCTAAGGAGGCCTGTAGGCTGCTGGGACTAAGTTTGCCGGCAAG  
GACCAAGTTGAAACAGCCTTGAAAGCTAGCTTTGAAACTTGACGCTATGGCTGGGTGGGAGA  
TGGATTTCGTGGTCATCTCTAGGATTAGCCCAAACCCCAAGTGTGGGAAAAATGGGGTGGGTG  
TCCTGATTTGGAAGGTTCCAGTGAGCCGACAGTTTGACAGCCTATTGTTACAACCTCATCTGAT  
ACTTGGAATAACTCGTGCATTCCAGAAATTATCACCACCAAAGATCCCATATTCAACACTCA  
AACTGCAACACAAACAACAGAATTTATTGTCAAGTGACAGTACCTACTCGGTGGCATCCCCTT  
ACTCTACAATACTGCCCTACTACTACTCCTCCTGCTCCAGCTTCCACTTCTATTCCACGG  
AGAAAAAAATTGATTTGTGTACAGAAAGTTTTTATGGAACTAGCACCATGTCTACAGAAAC  
TGAACCATTTGTTGAAAATAAAGCAGCATTCAAGAATGAAGCTGCTGGGTTTGAGAGGTGTCC  
CCACGGCTCTGCTAGTGCTTGCTCTCCTCTTCTTTGGTGCTGCAGCTGGTCTTGATTTTGC  
TATGTCAAAGGTATGTGAAGGCCTTCCCTTTTACAAACAAGAATCAGCAGAAGGAAATGAT  
CGAAACCAAAGTAGTAAAGGAGGAGAAGGCCAATGATAGCAACCCTAATGAGGAATCAAAGA  
AAACTGATAAAAACCCAGAAGAGTCCAAGAGTCCAAGCAAACTACCGTGCAGTGCCTGGAA  
GCTGAAGTTTAGATGAGACAGAAATGAGGAGACACACCTGAGGCTGGTTTCTTTCATGCTCC  
TTACCCTGCCCCAGCTGGGGAAATCAAAGGGCCAAAGAACCAGAAAGAAAGTCCACCCTT  
GGTTCCTAACTGGAATCAGCTCAGGACTGCCATTGGACTATGGAGTGCACCAAAGAGAATGC  
CCTTCTCCTTATTGTAAACCCTGTCTGGATCCTATCCTCCTACCTCAAAGCTTCCCACGGCC  
TTTCTAGCCTGGCTATGTCTTAATAATATCCCCTGAGGAGAAAGGAGTTTGTCAAAGTGCAA  
GGACCTAAACATCTCATCAGTATCCAGTGGTAAAAAGGCCTCCTGGCTGTCTGAGGCTAGG  
TGGGTTGAAAGCCAAGGAGTCACTGAGACCAAGGCTTCTCTACTGATTCCGCAGCTCAGAC  
CCTTTCTTTCAGCTCTGAAAGAGAAACACGATATCCACCTGACATGTCTTCTGAGCCCGGTA  
AGAGCAAAAGAATGGCAGAAAGTTTAGCCCCTGAAAGCCATGGAGATTCTCATAACTTGAG  
ACCTAATCTCTGTAAAGCTAAAATAAAGAAATAGAACAGGCTGAGGATACGACAGTACACT  
GTCAGCAGGGACTGTAAACACAGACAGGGTCAAAGTGTTTTCTCTGAACACATTGAGTTGGA  
ATCACTGTTTAGAACACACACACTTACTTTTTCTGGTCTCTACCCTGCTGATATTTTCTCT  
AGGAAATATACTTTTACAAGTAACAAAAATAAAACTCTTATAAATTTCTATTTTATCTGA  
GTTACAGAAATGATTACTAAGGAAGATTACTCAGTAATTTGTTTAAAAAGTAATAAAATTCA  
ACAAACATTTGCTGAATAGCTACTATATGTCAAGTGCTGTGCAAGGTATTACACTCTGTAAT  
TGAATATTATTCTCAAAAAATTGCACATAGTAGAACGCTATCTGGGAAGCTATTTTTCAT  
GTTTTGATATTTCTAGCTTATCTACTTCCAACTAATTTTATTTTGTGCTGAGACTAATCTT  
ATTCAATTTTCTCTAATATGGCAACCATTTATAACCTTAATTTATTATTAAACATACCTAAGAAG  
TACATTGTTACCTCTATATACCAAAGCACATTTTAAAAGTGCCATTAAACAAATGTATCACTA  
GCCCTCCTTTTTTCCAACAAGGAGTGAAGATGCAGAAATATTTGTGACAAAAAATTAA  
AGCATTTAGAAAACCTT

Sequence 73

[illegible]

amino acids 23-29, 65-71, 234-240, 235-239, 249-255, 253-259

## FIGURE 75

AGATGCGGTCTTGGCACCTCTAATTGCTCTCGTGATTTCGGTGCCGCGACTTTCACGATGG  
CTCGCCCAACCTTACTACCTTCTGTGCGCCCTGCTCTCTGCTGCCTTCCTACTCGTGAGGAA  
ACTGCCGCGCTCTGCCACGGTCTGCCCACCCAACGCGAAGACGGTAACCCGTGTGACTTTG  
ACTGGAGAGAAGTGGAGATCCTGATGTTTCTCAGTGCCATTGTGATGATGAAGAACCGCAGA  
TCCATCACTGTGGAGCAACATATAGGCAACATTTTCATGTTTAGTAAAGTGGCCAACACAAT  
TCTTTTCTTCCGCTTGGATATTCGCATGGGCCTACTTTACATCACACTCTGCATAGTGTTC  
TGATGACGTGCAAACCCCCCTATATATGGGGCCCTGAGTATATCAAGTACTTCAATGATAAA  
ACCATTGATGAGGAACTAGAACGGGACAAGAGGGTCACTTGGATTGTGGAGTTCTTTGCCAA  
TTGGTCTAATGACTGCCAATCATTTGCCCTATCTATGCTGACCTCTCCCTTAAATACAAC  
GTACAGGGCTAAATTTTGGGAAGGTGGATGTTGGACGCTATACTGATGTTAGTACGCGGTAC  
AAAGTGAGCACATCACCCCTCACCAAGCAACTCCCTACCCTGATCCTGTTCCAAGGTGGCAA  
GGAGGCAATGCGGCGGCCACAGATTGACAAGAAAGGACGGGCTGTCTCATGGACCTTCTCTG  
AGGAGAAATGTGATCCGAGAATTTAACTTAAATGAGCTATACCAGCGGGCCAAGAACTATCA  
AAGGCTGGAGACAATATCCCTGAGGAGCAGCCTGTGGCTTCAACCCCCACCACAGTGTGAGA  
TGGGGAAAACAAGAAGGATAAATAAGATCCTCACTTTGGCAGTGCTTCCTCTCCTGTCAATT  
CCAGGCTCTTTCCATAACCACAAGCCTGAGGCTGCAGCCTTTNATTNATGTTTTCCCTTTGG  
CTGNGACTGGNTGGGGCAGCATGCAGCTTCTGATTTTAAAGAGGCATCTAGGGAATTGTCAG  
GCACCCTACAGGAAGGCCTGCCATGCTGTGGCCAACCTGTTTCACTGGAGCAAGAAAGAGATC  
TCATAGGACGGAGGGGGAAATGGTTTCCCTCCAAGCTTGGGTCAGTGTGTTAACTGCTTATC  
AGCTATTCAGACATCTCCATGGTTTCTCCATGAAACTCTGTGGTTTCATCATTCCTTCTTAG  
TTGACCTGCACAGCTTGGTTAGACCTAGATTTAACCCTAAGGTAAGATGCTGGGGTATAGAA  
CGCTAAGAATTTTCCCCCAAGGACTCTTGCTTCCTTAAGCCCTTCTGGCTTCGTTTATGGTC  
TTCATTAAGAGTATAAGCCTAACTTTGTGCTAGTCCTAAGGAGAAACCTTTAACCACAAAG  
TTTTTATCATTGAAGACAATATTGAACAACCCCTATTTTGTGGGGATTGAGAAGGGGTGAA  
TAGAGGCTTGAGACTTTCCTTTGTGTGGTAGGACTTGGAGGAGAAATCCCCTGGACTTTCAC  
TAACCCTCTGACATACTCCCCACACCCAGTTGATGGCTTTCCGTAATAAAAAGATTGGGATT  
TCCTTTTG

## **FIGURE 76**

MAVLAPLIALVYSVPRLSRWLAQPYLLSALLSAAFLLVKLPPLCHGLPTQREDGNPCDFD  
WREVEILMFLSAIVMMKNRRSITVEQHIGNIFMFSKVANTILFFRLDIRMGLLYITLCIVFL  
MTCKPPLYMGPEYIKYFNDKTIDEELERDKRVTWIVEFFANWSNDCQSFAPIYADLSLKYNC  
TGLNFGKVDVGRYTDVSTRYKVSTSPLTKQLPTLILFQGGKEAMRRPQIDKKGRAVSWTFSE  
ENVIREFNLNELYQRAKKLSKAGDNIPEEQPVASTPTTVSDGENKKDK

### **Signal sequence:**

amino acids 1-48

### **Transmembrane domain:**

amino acids 111-125

### **N-glycosylation site.**

amino acids 165-169, 185-189

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 154-158, 265-269

### **Casein kinase II phosphorylation site.**

amino acids 51-55, 145-149, 245-249, 286-290, 288-292

### **N-myristoylation site.**

amino acids 188-194, 225-231

### **Myb DNA-binding domain repeat signature 1.**

amino acids 244-253

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



## FIGURE 77

GGACAGCTCGCGGCCCCGAGAGCTCTAGCCGTCGAGGAGCTGCCTGGGGACGTTTGCCCTG  
GGGCCCCAGCCTGGCCCCGGGTCAACCCTGGCATGAGGAGATGGGCCTGTTGCTCCTGGTCCCA  
TTGCTCCTGCTGCCCGGCTCCTACGGACTGCCCTTCTACAACGGCTTCTACTACTCCAACAG  
CGCCAACGACCAGAACCTAGGCAACGGTCATGGCAAAGACCTCCTTAATGGAGTGAAGCTGG  
TGGTGGAGACACCCGAGGAGACCCTGTTACCTACCAAGGGGCCAGTGTGATCCTGCCCTGC  
CGCTACCGCTACGAGCCGGCCCTGGTCTCCCCGCGGCGTGTGCGTGTCAAATGGTGGAAGCT  
GTCGGAGAACGGGGCCCCAGAGAAGGACGTGCTGGTGGCCATCGGGCTGAGGCACCGCTCCT  
TTGGGGACTACCAAGGCCGCGTGCACCTGCGGCAGGACAAAGAGCATGACGTCTCGCTGGAG  
ATCCAGGATCTGCGGCTGGAGGACTATGGGCGTTACCGCTGTGAGGTCAATTGACGGGCTGGA  
GGATGAAAGCGGTCTGGTGGAGCTGGAGCTGCGGGGTGTGGTCTTTCTTACCAGTCCCCCA  
ACGGGCGCTACCAGTTCAACTTCCACGAGGGGCCAGCAGGTCTGTGCAGAGCAGGCTGCGGTG  
GTGGCCTCCTTTGAGCAGCTCTTCCGGGCCTGGGAGGAGGGCCTGGACTGGTGCAACGCGGG  
CTGGCTGCAGGATGCTACGGTGCAGTACCCCATCATGTTGCCCGGCAGCCCTGCGGTGGCC  
CAGGCCTGGCACCTGGCGTGCGAAGCTACGGCCCCCGCCACCGCCGCTGCACCGCTATGAT  
GTATTCTGCTTCGCTACTGCCCTCAAGGGGCGGGTGTACTACCTGGAGCACCTGAGAAGCT  
GACGCTGACAGAGGCAAGGGAGGCCTGCCAGGAAGATGATGCCACGATCGCCAAGGTGGGAC  
AGCTCTTTGCCGCTGGAAGTTCCATGGCCTGGACCGCTGCGACGCTGGCTGGCTGGCAGAT  
GGCAGCGTCCGCTACCCCTGTGGTTCACCCGCATCCTAACTGTGGGGCCCCAGAGCCTGGGGT  
CCGAAGCTTTGGCTTCCCCGACCCGCAGAGCCGCTTGTACGGTGTTTACTGCTACCGCCAGC  
ACTAGGACCTGGGGCCCTCCCCTGCCGCATTCCCTCACTGGCTGTGTATTTATTGAGTGGTT  
CGTTTTCCCTTGTGGGTGGAGCCATTTTAACTGTTTTTATACTTCTCAATTTAAATTTTCT  
TTAAACATTTTTTTTACTATTTTTTTGTAAAGCAAACAGAACCCAATGCCTCCCTTTGCTCCTG  
GATGCCCCACTCCAGGAATCATGCTTGCTCCCCCTGGGCCATTTGCGGTTTTGTGGGCTTCTG  
GAGGGTTCCCCGCCATCCAGGCTGGTCTCCCTCCCTTAAGGAGGTTGGTGCCAGAGTGGGC  
GGTGGCCTGTCTAGAATGCCGCCGGGAGTCCGGGCATGGTGGGCACAGTTCTCCCTGCCCT  
CAGCCTGGGGGAAGAAGAGGGCCTCGGGGGCCTCCGAGCTGGGCTTTGGGCCTCTCCTGCC  
CACCTTACTTCTCTGTGAAGCCGCTGACCCAGTCTGCCCACTGAGGGGCTAGGGCTGGAA  
GCCAGTTCTAGGCTTCCAGGCGAAATCTGAGGGAAGGAAGAACTCCCCCTCCCCGTTCCCCCT  
TCCCCTCTCGGTTCCAAAGAATCTGTTTTGTTGTGTCATTTGTTTCTCCTGTTTCCCTGTGTGG  
GGAGGGGCCCTCAGGTGTGTGTACTTTGGACAATAAATGGTGCTATGACTGCCTTCCGCCAA  
AA  
AA

## FIGURE 78

MGLLLLVPLLLLPGSYGLPFYNGFYYSNSANDQNLGNHGHGKDLLNGVKLVVETPEETLFTYQ  
GASVILPCRYRYEPALVSPRRVRVKWWKLSENGAPEKDVLAIGLRHRSFGDYQGRVHLRQD  
KEHDVSLEIQDLRLLEDYGRYRCEVIDGLEDESGLVELELRGVVFPYQSPNGRYQFNFHEGQQ  
VCAEQAAVVASFEQLFRAWEEGLDWCNAGWLQDATVQYPIMLPRQPCGGPGLAPGVRSYGPR  
HRLHRYDVFCFATALKGRVYYLEHPEKLTLTEAREACQEDDATIAKVGQLFAAWKFHGLDR  
CDAGWLADGSVRYPVVHHPNCGPPEPGVRSFGFPDPQSRLYGVYCYRQH

### **Signal sequence:**

amino acids 1-17

### **Casein kinase II phosphorylation site.**

amino acids 29-33, 53-57, 111-115, 278-282

### **Tyrosine kinase phosphorylation site.**

amino acids 137-145

### **N-myristoylation site.**

amino acids 36-42, 184-190, 208-214, 237-243, 297-303, 307-313

peptide "chasing"

## FIGURE 79

GGAGAGCGGAGCGAAGCTGGATAACAGGGGACCGATGATGTGGCGACCATCAGTTCTGCTGC  
TTCTGTTGCTACTGAGGCACGGGGCCCAGGGGAAGCCATCCCCAGACGCAGGCCCTCATGGC  
CAGGGGAGGGTGCACCAGGCGGCCCCCTGAGCGACGCTCCCCATGATGACGCCACGGGAA  
CTTCCAGTACGACCATGAGGCTTTCTCTGGGACGGGAAGTGGCCAAGGAATTCGACCAACTCA  
CCCCAGAGGAAAGCCAGGCCCGTCTGGGGCGGATCGTGGACCGCATGGACCGCGCGGGGGAC  
GGCGACGGCTGGGTGTCGCTGGCCGAGCTTCGCGCGTGGATCGCGCACACGCAGCAGCGGCA  
CATACGGGACTCGGTGAGCGCGGCCTGGGACACGTACGACACGGACCGCGACGGGCGTGTGG  
GTTGGGAGGAGCTGCGCAACGCCACCTATGGCCACTACGCGCCCCGGTGAAGAATTTTCATGAC  
GTGGAGGATGCAGAGACCTACAAAAAGATGCTGGCTCGGGACGAGCGGCGTTTCCGGGTGGC  
CGACCAGGATGGGGACTCGATGGCCACTCGAGAGGAGCTGACAGCCTTCTCTGCACCCCGAGG  
AGTTCCCTCACATGCGGGACATCGTGATTGCTGAAACCCTGGAGGACCTGGACAGAAACAAA  
GATGGCTATGTCCAGGTGGAGGAGTACATCGCGGATCTGTACTCAGCCGAGCCTGGGGAGGA  
GGAGCCGGCGTGGGTGCAGACGGAGAGGCAGCAGTTCCGGGACTTCCGGGATCTGAACAAGG  
ATGGGCACCTGGATGGGAGTGAGGTGGGCCACTGGGTGCTGCCCCCTGCCCAGGACCAGCCC  
CTGGTGGAAGCCAACCACCTGCTGCACGAGAGCGACACGGACAAGGATGGGCGGCTGAGCAA  
AGCGGAAATCCTGGGTAAATTGGAACATGTTTGTGGGCAGTCAGGCCACCAACTATGGCGAGG  
ACCTGACCCGGCACCCACGATGAGCTGTGAGCACCGCGCACCTGCCACAGCCTCAGAGGCCCG  
CACAATGACCGGAGGAGGGGCGCTGTGGTCTGGCCCCCTCCCTGTCCAGGCCCCCGCAGGAG  
GCAGATGCAGTCCCAGGCATCCTCCTGCCCCCTGGGCTCTCAGGGACCCCCTGGGTCTGGCTTC  
TGTCCTGTGCACACCCCCAACCCAGGGAGGGGCTGTCTAGTCCCAGAGGATAAGCAATAC  
CTATTTCTGACTGAGTCTCCCAGCCCAGACCCAGGGACCCTTGGCCCCAAGCTCAGCTCTAA  
GAACCGCCCCAACCCCTCCAGCTCCAAATCTGAGCCTCCACCACATAGACTGAAACTCCCCT  
GGCCCCAGCCCTCTCCTGCCTGGCCTGGCCTGGGACACCTCCTCTCTGCCAGGAGGCAATAA  
AAGCCAGCGCCGGGACCTTGAAA  
AAAAAAAAAAAAAAAAA

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## **FIGURE 80**

MMWRPSVLLLLLLLLLRHGAQGKPSPDAGPHGQGRVHQAAPLSDA PHDDAHGNFQYDHEAFLGR  
EVAKEFDQLTP EESQARLGRI VDRMDRAGDGDGWVSLAELRAWIAHTQQRHIRDSVSAAWDT  
YDTRDRDGRVGWEELRNATYGHYAPGEEFHDVEDAETYKKMLARDERRFRVADQDGD SMATRE  
ELTAFLHP EEFPHMRDIVIAETLEDLDRNKDGYVQVEEYIADLYSAEPGEEEP AWVQTERQQ  
FRDFRDLNKDGHLDGSEVGHWVLP PAQDQPLVEANHLLHESD TDKDGRLSKAEILGNWNMFV  
GSQATNYGEDLTRHHDEL

### **Signal sequence:**

amino acids 1-20

### **N-glycosylation site.**

amino acids 140-144

### **Casein kinase II phosphorylation site.**

amino acids 72-76, 98-102, 127-131, 184-188, 208-212, 289-293,  
291-295, 298-302

### **N-myristoylation site.**

amino acids 263-269, 311-317

### **Endoplasmic reticulum targeting sequence.**

amino acids 325-330

## FIGURE 81

GGGGCCTTGCCTTCCGCACTCGGGCGCAGCCGGGTGGATCTCGAGCAGGTGCGGAGCCCCGG  
GCGGCGGGCGCGGGTGCAGGGGATCCCTGACGCCCTCTGTCCCTGTTTCTTTGTGCTCCCAG  
CCTGTCTGTCTGTCGTTTTTGGCGCCCCCGCCTCCCCGCGGTGCGGGGTTGCACACCGATCCTG  
GGCTTCGCTCGATTGCGCGCCGAGGCGCTCCAGACCTAGAGGGGCGCTGGCCTGGAGCAG  
CGGGTCGTCTGTGTCCTCTCTCCTCTGCGCCGCGCCCCGGGGATCCGAAGGGTGCGGGGCTCT  
GAGGAGGTGACGCGCGGGGCCTCCCGCACCTTGGCCTTGCCCGCATTCTCCCTCTCTCCCAG  
GTGTGAGCAGCCTATCAGTCACCAATGTCCGCGAGCCTGGATCCCGGCTCTCGGCCTCGGTGTG  
TGTCTGCTGCTGCTGCCGGGGCCCCGCGGGCAGCGAGGGAGCCGCTCCCATTTGCTATCACATG  
TTTTACCAGAGGCTTGACATCAGGAAAGAGAAAGCAGATGTCCTCTGCCAGGGGGCTGCC  
CTCTTGAGGAATTCTCTGTGTATGGGAACATAGTATATGCTTCTGTATCGAGCATATGTGGG  
GCTGCTGTCCACAGGGGAGTAATCAGCAACTCAGGGGGACCTGTACGAGTCTATAGCCTACC  
TGGTCGAGAAAATATTCTCAGTAGATGCCAATGGCATCCAGTCTCAAATGCTTTCTAGAT  
GGTCTGCTTCTTTCACAGTAACTAAAGGCAAAAGTAGTACACAGGAGGCCACAGGACAAGCA  
GTGTCCACAGCACATCCACCAACAGGTAAACGACTAAAGAAAACACCCGAGAAGAAAATGG  
CAATAAAGATTGTAAAGCAGACATTGCATTTCTGATTGATGGAAGCTTTAATATTGGGCAGC  
GCCGATTTAATTTACAGAAGAATTTTGTGGGAAAAGTGGCTCTAATGTTGGGAATTGGAACA  
GAAGGACCACATGTGGGCCTTGTTCAGCCAGTGAACATCCCAAATAGAATTTTACTTGAA  
AACTTTTACATCAGCCAAAGATGTTTTGTTTGCCATAAAGGAAGTAGGTTTCAGAGGGGGTA  
ATTCCAATACAGGAAAAGCCTTGAAGCATACTGCTCAGAAATTCTTCACGGTAGATGCTGGA  
GTAAGAAAAGGGATCCCCAAAGTGGTGGTGGTATTTATTGATGGTTGGCCTTCTGATGACAT  
CGAGGAAGCAGGCATTGTGGCCAGAGAGTTTGGTGTCAATGTATTTATAGTTTCTGTGGCCA  
AGCCTATCCCTGAAGAACTGGGGATGGTTCAGGATGTCACATTTGTTGACAAGGCTGTCTGT  
CGGAATAATGGCTTCTTCTTACCACATGCCCAACTGGTTTGGCACCAAAAATACGTAAA  
GCCTCTGGTACAGAAGCTGTGCACTCATGAACAAATGATGTGCAGCAAGACCTGTTATAACT  
CAGTGAACATTGCCTTTCTAATTGATGGCTCCAGCAGTGTGGAGATAGCAATTTCCGCCTC  
ATGCTTGAATTTGTTTCCAACATAGCCAAGACTTTTGAAATCTCGGACATTGGTGCCAAGAT  
AGCTGCTGTACAGTTTACTTATGATCAGCGCACGGAGTTCAGTTTCACTGACTATAGCACCA  
AAGAGAATGTCCTAGCTGTCATCAGAAACATCCGCTATATGAGTGGTGGAACAGCTACTGGT  
GATGCCATTTCTTCACTGTTAGAAATGTGTTTGGCCCTATAAGGGAGAGCCCCAACAAGAA  
CTTCTAGTAATTGTCACAGATGGGCAGTCTATGATGATGTCCAAGGCCCTGCAGCTGCTG  
CACATGATGCAGGAATCACTATCTTCTCTGTTGGTGTGGCTTGGGCACCTCTGGATGACCTG  
AAAGATATGGCTTCTAAACCGAAGGAGTCTCACGCTTTCTTACAAGAGAGTTTACAGGATT  
AGAACCAATTGTTTCTGATGTCATCAGAGGCATTTGTAGAGATTTCTTAGAATCCCAGCAAT  
AATGGTAACATTTTGACAACTGAAAGAAAAAGTACAAGGGGATCCAGTGTGTAAATTGTATT  
CTCATAATACTGAAATGCTTTAGCATACTAGAATCAGATACAAAATATTAAGTATGTCAAC  
AGCCATTTAGGCAAATAAGCACTCCTTTAAAGCCGCTGCCTTCTGGTTACAATTTACAGTGT  
ACTTTGTTAAAAACACTGCTGAGGCTTCATAATCATGGCTCTTAGAAACTCAGGAAAGAGGA  
GATAATGTGGATTAAAACCTTAAGAGTTCTAACCATGCCTACTAAATGTACAGATATGCAAA  
TTCCATAGCTCAATAAAAGAATCTGATACTTAGACCAAAAAAAAAA

## **FIGURE 82**

MSAAWIPALGLGVCLLLLPGPAGSEGAAPIAITCFTRGLDIRKEKADVLCPPGGCPLEEF SVY  
GNIVYASVSSICGAAVHRGVISNSGGPVRVYSLPGRENYSSVDANGIQSQMLSRWSASF T VT  
KGKSSTQEATGQAVSTAHPPTGKRLKKTPEKKTGNKDKADIAFLIDGSFNIGQRRFNLQKN  
FVGKVALMLGIGTEGPHVGLVQASEHPKIEFYLNFTSAKDVLF AIKEVGFRGGNSNTGKAL  
KHTAQKFFTVDAGVRKGIPKVVVVFIDGWPSDDIEEAGIVAREFGVNVFIVSVAKPIPEELG  
MVQDVT FVDKAVCRNNGFFSYHMPNWF GTTKYVKPLVQKLCTHEQMMCSKTCYNSVNIAFLI  
DGSSSVGDSNFRLMLEFVSNIAKTFEISDIGAKIAAVQFTYDQRTEFSFTDYSTKENVLAVI  
RNIRYMSGGTATGDAISFTVRNVFGPIRESPNKNFLVIVTDGQSYDDVQGPAAAAH DAGITI  
FSVGVAWAPLDDLKDMASKPKESHAF FTREFTGLEPIVSDVIRGICRDFLESQQ

### **Signal sequence:**

amino acids 1-24

### **N-glycosylation site.**

amino acids 100-104, 221-225

### **Casein kinase II phosphorylation site.**

amino acids 102-106, 129-133, 224-228, 316-320, 377-381, 420-424,  
425-429, 478-482, 528-532

### **N-myristoylation site.**

amino acids 10-16, 23-29, 81-87, 135-141, 158-164, 205-211,  
239-245, 240-246, 261-267, 403-409, 442-448, 443-449

### **Amidation site.**

amino acids 145-149

## FIGURE 83

CGCCGCGCTCCCGCACCCGCGGCCCGCCACCGCGCGCTCCCGCATCTGCACCCGCAGCCC  
GGCGGCCTCCCGGCGGGAGCGAGCAGATCCAGTCCGGGCCGCGAGCGCAACTCGGTCCAGTCG  
GGGCGGCGGCTGCGGGCGCAGAGCGGAGATGCAGCGGCTTGGGGCCACCCTGCTGTGCCTGC  
TGCTGGCGGCGGCGGTCCCCACGGCCCCCGCGCCGCTCCGACGGCGACCTCGGCTCCAGTC  
AAGCCCGGCCCGGCTCTCAGCTACCCGCAGGAGGAGGCCACCCTCAATGAGATGTTCCGCGA  
GGTTGAGGAAGCTGCTGCTAAAGCATCATCAGAAGTGAACCTGGCAAACCTTACCTCCAGCTAT  
CACAATGAGACCAACACAGACACGAAGGTTGGAAATAATACCATCCATGTGCACCGAGAAAT  
TCACAAGATAACCAACAACCAGACTGGACAAATGGTCTTTTTCAGAGACAGTTATCACATCTG  
TGGGAGACGAAGAAGGCAGAAGGAGCCACGAGTGCATCATCGACGAGGACTGTGGGGCCAGC  
ATGTACTGCCAGTTTGCCAGCTTCCAGTACACCTGCCAGCCATGCCGGGGCCAGAGGATGCT  
CTGCACCCGGGACAGTGAGTGCTGTGGAGACCAGCTGTGTGTCTGGGGTCACTGCACCAAAA  
TGGCCACCAGGGGCAGCAATGGGACCATCTGTGACAACCAGAGGGACTGCCAGCCGGGGCTG  
TGCTGTGCCTTCCAGAGAGGCCTGCTGTTCCCTGTGTGCACACCCCTGCCCGTGGAGGGCGA  
GCTTTGCCATGACCCCGCCAGCCGGCTTCTGGACCTCATCACCTGGGAGCTAGAGCCTGATG  
GAGCCTTGAGACCGATGCCCTTGTGCCAGTGGCCTCCTCTGCCAGCCCCACAGCCACAGCCTG  
GTGTATGTGTGCAAGCCGACCTTCGTGGGGAGCCGTGACCAAGATGGGGAGATCCTGCTGCC  
CAGAGAGGTCCCCGATGAGTATGAAGTTGGCAGCTTCATGGAGGAGGTGCGCCAGGAGCTGG  
AGGACCTGGAGAGGAGCCTGACTGAAGAGATGGCGCTGGGGGAGCCTGCGGCTGCCGCCGCT  
GCACTGCTGGGAGGGGAAGAGATTTAGATCTGGACCAGGCTGTGGGTAGATGTGCAATAGAA  
ATAGCTAATTTATTTCCCCAGGTGTGTGCTTTAGGCGTGGGCTGACCAGGCTTCTTCCTACA  
TCTTCTTCCAGTAAGTTTCCCTCTGGCTTGACAGCATGAGGTGTTGTGCATTTGTTTCAGC  
TCCCCAGGCTGTTCTCCAGGCTTCCAGCTGCTGGTGGGAGAGTCAGGCAGGGTTAAAC  
TGCAGGAGCAGTTTGCCACCCCTGTCCAGATTATTGGCTGCTTTGCCTCTACCAGTTGGCAG  
ACAGCCGTTTGTCTACATGGCTTTGATAATTGTTTGGGGGAGGAGATGGAAACAATGTGG  
AGTCTCCCTCTGATTGTTTTGGGGAAATGTGGAGAAGAGTGCCCTGCTTTGCAAAACATCAA  
CCTGGCAAAAATGCAACAAATGAATTTTCCACGCAGTTCTTTCCATGGGCATAGGTAAGCTG  
TGCCTTCAGCTGTTGCAGATGAAATGTTCTGTTACCCCTGCATTACATGTGTTTATTTCATCC  
AGCAGTGTTGCTCAGCTCCTACCTCTGTGCCAGGGCAGCATTTTTCATATCCAAGATCAATTC  
CCTCTCTCAGCACAGCCTGGGGAGGGGGTCATTGTTCTCCTCGTCCATCAGGGATCTCAGAG  
GCTCAGAGACTGCAAGCTGCTTGCCCAAGTCACACAGCTAGTGAAGACCAGAGCAGTTTCAT  
CTGGTTGTGACTCTAAGCTCAGTGCTCTCTCCACTACCCACACCAGCCTTGGTGCCACCAA  
AAGTGCTCCCCAAAAGGAAGGAGAATGGGATTTTTCTTGAGGCATGCACATCTGGAATTAAG  
GTCAAACATAATTCTCACATCCCTCTAAAAGTAACTACTGTTAGGAACAGCAGTGTTCTCAC  
AGTGTGGGGCAGCCGTCCTTCTAATGAAGACAATGATATTGACACTGTCCCTCTTTGGCAGT  
TGCATTAGTAACCTTGAAAGGTATATGACTGAGCGTAGCATACAGGTTAACCTGCAGAAACA  
GTAAGACAATTATCAACCACGTGGAGAAAAATCAAACCGAGCAGGGCTGTGTGAAACATGGTT  
GTAATATGCGACTGCGAACACTGAACTCTACGCCACTCCACAAATGATGTTTTTCAGGTGTCA  
TGGACTGTTGCCACCATGTATTTCATCCAGAGTTCTTAAAGTTTAAAGTTGCACATGATTGTA  
TAAGCATGCTTTCTTTGAGTTTTTAAATTATGTATAAACATAAGTTGCATTTAGAAATCAAGC  
ATAAATCACTTCAACTGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

CGCCGCGCTCCCGCACCCGCGGCCCGCCACCGCGCGCTCCCGCATCTGCACCCGCAGCCC

## FIGURE 84

MQRLGATLLCLLLAAVPTAPAPAPTATSAPVKPGPALSYPQEEATLNEMFREVEELMEDTQ  
HKLRSAVEEMEAEEAAAKASSEVNLANLPPSYHNETNTDTKVGNNTIHVHREIHKITNNQTG  
QMVFSETVITSVGDEEGRRSHECIIDEDCGPSMYCQFASFQYTCQPCRQGRMLCTRDSECCG  
DQLCVWGHCTKMATRGSNGTICDNQRDCQPGLCCAFQRGLLFPVCTPLPVEGELCHDPASRL  
LDLITWELEPDGALDRCPCASGLLCQPHSHSLVYVCKPTFVGSRDQDGEILLPREVPDEYEV  
GSFMEEVRQELEDLERSLTEEMALGEPAAAAAALLGGEI

### **Signal sequence:**

amino acids 1-19

### **N-glycosylation site.**

amino acids 96-100, 106-110, 121-125, 204-208

### **Casein kinase II phosphorylation site.**

amino acids 46-50, 67-71, 98-102, 135-139, 206-210, 312-316,  
327-331

### **N-myristoylation site.**

amino acids 202-208, 217-223

### **Amidation site.**

amino acids 140-144

6443000



[illegible]

AAAAA

## **FIGURE 86**

MRLLVAPLLLAWVAGATATVPVVPWHVPCPPQCACQIRPWYTPRSSYREATTVDCNDLFLTA  
VPPALPAGTQTLTLLQSNSIVRVDQSELGYLANLTDLDSQNSFSDARDCDFHALPQLLSLHL  
EENQLTRLEDHSFAGLASLQELYLNHNQLYRIAPRAFSGLSNLLRLHLNSNLLRAIDSRWFE  
MLPNLEILMIGGNKVDAILD MNFRPLANLRSLVLAGMNLREISDYALEGLQSLESLSFYDNQ  
LARVPRRALEQVPGLKFLDLNKNPLQRVGPGDFANMLHLKELGLNNMEELV SIDKFALVNLP  
ELTKLDITNNPRLSFIHPRAFHHLPQMETLMLNNNALSALHQQTVESLPNLQEVGLHGNPIR  
CDCVIRWANATGTRVRFIEPQSTLCAEPPDLQRLPVREVPPFREMTDHCLPLISPRSFPFSLQ  
VASGESMVLHCRALAEPEPEIYWVTPAGLRLTPAHAGRRYRVYPEGTLELRRVTAAEEAGLYT  
CVAQNLVGADTKTVSVVGRALLQPGRDEGQGLELRVQETHPYHILLSWVTPPNTVSTNLTW  
SSASSLRGQGATALARLPRGTHSYNITRLLQATEYWACLQVAFADAHTQLACVWARTKEATS  
CHRALGDRPGLIAILALAVLLLAAGLAHLGTGQPRKGVGRRPLPPAWAFWGWSAPSVRVV  
SAPLVLPWNPGRKLPRSSEGETLLPPLSQNS

### **Signal sequence:**

amino acids 1-18

### **Transmembrane domain:**

amino acids 629-648

### **N-glycosylation site.**

amino acids 94-98, 381-385, 555-559, 583-587

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 485-489

### **Casein kinase II phosphorylation site.**

amino acids 46-50, 51-55, 96-100, 104-108, 130-134, 142-146,  
243-247, 313-317, 488-492, 700-704

### **Tyrosine kinase phosphorylation site.**

amino acids 532-540

### **N-myristoylation site.**

amino acids 15-21, 493-499, 566-572

### **Amidation site.**

amino acids 470-474, 660-664, 692-696

## FIGURE 87

GCAAGCCAAGGCGCTGTTTGAGAAGGTGAAGAAGTTCCGGACCCATGTGGAGGAGGGGGACATTGTGTACCGCCT  
CTACATGCGGCAGACCATCATCAAGGTGATCAAGTTCATCCTCATCATCTGCTACACCGTCTACTACGTGCACAA  
CATCAAGTTCGACGTGGACTGCACCGTGGACATTGAGAGCCTGACGGGCTACCGCACCTACCGCTGTGCCCACCC  
CCTGGCCACACTCTTCAAGATCCTGGCGTCTTCTACATCAGCCTAGTCATCTTCTACGGCCTCATCTGCATGTA  
CACACTGTGGTGGATGCTACGGCGCTCCCTCAAGAAGTACTCGTTTGAGTCGATCCGTGAGGAGAGCAGCTACAG  
CGACATCCCCGACGTCAAGAACGACTTCGCCCTTCATGCTGCACCTCATTGACCAATACGACCCGCTCTACTCCAA  
GCGCTTCGCCGCTCTTCTGTGCGGAGGTGAGTGAGAACAAGCTGCGGCAGCTGAACCTCAACAACGAGTGGACGCT  
GGACAAGCTCCGGCAGCGGCTCACCAAGAACGCGCAGGACAAGCTGGAGCTGCACCTGTTTCATGCTCAGTGGCAT  
CCCTGACACTGTGTTTGACCTGGTGGAGCTGGAGGTCTCAAGCTGGAGCTGATCCCCGACGTGACCATCCCCGCC  
CAGCATTGCCCAGCTCAGGGCCTCAAGGAGCTGTGGCTCTACCACACAGCGGCCAAGATTGAAGCGCCTGCGCT  
GGCCTTCTGCGCGAGAACCTGCGGGCGCTGCACATCAAGTTCACCGACATCAAGGAGATCCCGCTGTGGATCTA  
TAGCCTGAAGACACTGGAGGAGCTGCACCTGACGGGCAACCTGAGCGCGGAGAACAACCGCTACATCGTCATCGA  
CGGGCTGCGGGAGCTCAAACGCCTCAAGGTGCTGCGGCTCAAGAGCAACCTAAGCAAGCTGCCACAGGTGGTTCAC  
AGATGTGGGCGTGCACCTGCAGAAGCTGTCCATCAACAATGAGGGCACCAAGCTCATCGTCTCAACAGCCTCAA  
GAAGATGGCGAACCTGACTGAGCTGGAGCTGATCCGCTGCGACCTGGAGCGCATCCCCACTCCATCTTCAGCCT  
CCACAACCTGCAGGAGATTGACCTCAAGGACAACAACCTCAAGACCATCGAGGAGATCATCAGCTTCCAGCACCT  
GCACCGCCTCACCTGCCTTAAGCTGTGGTACAACCACATCGCCTACATCCCCATCCAGATCGGCAACCTCACCAA  
CCTGGAGCGCCTCTACCTGAACCGCAACAAGATCGAGAAGATCCCCACCCAGCTCTTCTACTGCCGCAAGCTCGG  
CTACCTGGACCTCAGCCACAACAACCTGACCTTCTCCCTGCCGACATCGGCCTCCTGCAGAACCTCCAGAACCT  
AGCCATCACGGCCAACCGGATCGAGACGCTCCCTCCGGAGCTCTTCCAGTGCCGGAAGCTGCGGGCCCTGCACCT  
GGGCAACAACGTGCTGCAGTCACTGCCCTCCAGGTGGGCGAGCTGACCAACCTGACGCAGATCGAGCTGCGGGG  
CAACCGGCTGGAGTGCCTGCCTGTGGAGCTGGGCGAGTGCCCACTGCTCAAGCGCAGCGGCTTGGTGGTGGAGGA  
GGACCTGTTCAACACACTGCCACCCGAGGTGAAGGAGCGGCTGTGGAGGGCTGACAAGGAGCAGGCTGAGCGGAG  
GCCGGCCAGCACAGCAAGCAGCAGGACCGCTGCCAGTCTCAGGCCCGGAGGGGAGGGCCTAGCTTCTCCAG  
AACTCCCGGACAGCCAGGACAGCCTCGCGGCTGGGCGAGGAGCCTGGGGCCGCTTGTGAGTCAGGCGAGCGAGA  
GGACAGTATCTGTGGGGCTGGCCCCCTTTTCTCCCTCTGAGACTCACGTCCCCCAGGGCAAGTGCTTGTGGAGGAG  
AGCAAGTCTCAAGAGCGCAGTATTTGGATAATCAGGTCTCCTCCCTGGAGGCCAGCTCTGCCCCAGGGGCTGAG  
CTGCCACCAGAGGTCTGGGACCTCACTTTAGTTCTTGGTATTTATTTTTCTCCATCTCCCACCTCCTTCATCC  
AGATAACTTATACATTCCCAAGAAAGTTCAGCCCAGATGGAAGGTGTTTCAAGGAAAGGTGGGCTGCCTTTTCCCC  
TTGTCTTTATTTAGCGATGCCGCCGGGCATTTAACACCCACCTGGACTTCAGCAGAGTGGTCCGGGGCGAACCAG  
CCATGGGACGGTCAACCAGCAGTGCCGGGCTGGGCTCTGCGGTGCGGTCCACGGGAGAGCAGGCCTCCAGCTGGA  
AAGGCCAGGCCTGGAGCTTGCCTCTTCAGTTTTTGTGGCAGTTTTAGTTTTTTTGTTTTTTTTTTTTAAATCAAA  
AAACAATTTTTTTTTTAAAAAAGCTTTGAAAAATGGATGGTTTTGGGTATTAAAAAGAAAAAAGCTTAAAAA  
AAAAGACACTAACGGCCAGTGAGTTGGAGTCTCAGGGCAGGGTGGCAGTTTTCCCTTGAGCAAGCAGCCAGACGT  
TGAAGTGTGTTTCTTTCCCTGGGCGCAGGGTGCAGGGTGTCTTCCGATCTGGTGTGACCTTGGTCCAGGAGTT  
CTATTTGTTTCCCTGGGGAGGGAGGTTTTTTTTGTTTTGTTTTTTGGGTTTTTTTTGGTGTCTTGTTTCTTTCTCCTC  
ATGTGTCTTGGCAGGCACTCATTTCTGTGGCTGTGCGCCAGAGGGAATGTTCTGGAGCTGCCAAGGAGGGAGGAG  
ACTCGGGTTGGCTAATCCCCGATGAACGGTGCTCCATTTCGCACCTCCCCTCCTCGTGCCTGCCCTGCCCTCTCCA  
CGCACAGTGTTAAGGAGCCAAGAGGAGCCACTTCGCCCAGACTTTGTTTTCCCCACCTCCTGCGGCATGGGTGTGT  
CCAGTGCCACCGCTGGCCTCCGCTGCTTCCATCAGCCCTGTGCGCACCTGGTCTTTCATGAAGAGCAGACACTTA  
GAGGCTGGTCCGGAATGGGGAGGTGCCCCCTGGGAGGGCAGGCGTTGGTTCCAAGCCGGTTCCCGTCCCCTGGCGC  
CTGGAGTGACACAGCCAGTCGGCACCTGGTGGCTGGAAGCCAACCTGCTTTAGATCACTCGGGTCCCCACCTT  
AGAAGGGTCCCCGCCTTAGATCAATCAGTGGACACTAAGGCACGTTTTAGAGTCTCTTGTCTTAATGATTATGT  
CCATCCGTCTGTCCGTCCATTTGTGTTTTCTGCGTCTGTGATTGGATATAATCCTCAGAAATAATGCACACTAG  
CCTCTGACAACCATGAAGCAAAAATCCGTTACATGTGGGTCTGAACCTGTAGACTCGGTACAGTATCAAATAAA  
ATCTATAACAGAAAAA

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2015. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## **FIGURE 88**

MRQTIKVIKFIILICYTVYYVHNIKFDVDCTVDIESLTGYRTYRCAHPLATLFKILASFYI  
SLVIFYGLICMYTLWWMLRRSLKKYSFESIREESSYSDIPDVKNDFAFMLHLIDQYDPLYSK  
RFAVFLSEVSENKLRQLNLNNEWTLDKLRQLTKNAQDKLELHLFMLSIGIPDTVFDLVELEV  
LKLELIPDVTIPPSIAQLTGLKELWLYHTAAKIEAPALAFLENLRLALHIKFTDIKEIPLWI  
YSLKTEELHLTGNLNLSAENNRYIVIDGLRELKRLKVLRLKSNLSKLPQVVTDVGVHLQKLSI  
NNEGTKLIVLNSLKKMANLTELELIRCDLERIPHSIFSLHNLQEIDLKDNNLKTIEEIIISFQ  
HLHRLTCLKLWYNHIAIYIPIQIGNLTNLERLYLNRNKIEKIPTQLFYCRKLRYDLSHNNLT  
FLPADIGLLQNLQNLAITANRIETLPPPELFQCRKLRLALHLGNNVLQSLPSRVGELTNLTQIE  
LRGNRLECLPVELGECPLLKRSGLVVEEDLFNTLPPEVKERLWRADKEQA

### **Transmembrane domain:**

amino acids 51-75 (type II)

### **N-glycosylation site.**

amino acids 262-266, 290-294, 328-332, 396-400, 432-436, 491-495

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 85-89

### **Casein kinase II phosphorylation site.**

amino acids 91-95, 97-101, 177-181, 253-257, 330-334, 364-368,  
398-402, 493-497

### **N-myristoylation site.**

amino acids 173-179, 261-267, 395-401, 441-447

## FIGURE 89

GCCTGTTGCTGATGCTGCCGTGCGGTACTTGTCATGGAGCTGGCACTGCGGCGCTCTCCCGT  
CCCGCGGTGGTTGCTGCTGCTGCCGTGCTGCTGGGCCTGAACGCAGGAGCTGTCATTGACT  
GGCCCACAGAGGAGGGCAAGGAAGTATGGGATTATGTGACGGTCCGCAAGGATGCCTACATG  
TTCTGGTGGCTCTATTATGCCACCAACTCCTGCAAGAACTTCTCAGAACTGCCCCCTGGTCAT  
GTGGCTTCAGGGCGGTCCAGGCGGTTCTAGCACTGGATTTTGAAACTTTGAGGAAATTGGGC  
CCCTTGACAGTGATCTCAAACCACGGAAAACCACCTGGCTCCAGGCTGCCAGTCTCCTATTT  
GTGGATAATCCCGTGGGCACTGGGTTCACTTATGTGAATGGTAGTGGTGCCTATGCCAAGGA  
CCTGGCTATGGTGGCTTCAGACATGATGGTTCTCCTGAAGACCTTCTTCAGTTGCCACAAAG  
AATTCAGACAGTTCATTCTACATTTTCTCAGAGTCCTATGGAGGAAAAATGGCAGCTGGC  
ATTGGTCTAGAGCTTTATAAGGCCATTTCAGCGAGGGACCATCAAGTGCAACTTTGCGGGGGT  
TGCCTTGGGTGATTCTGATCTCCCCTGTTGATTTCGGTGCTCTCCTGGGGACCTTACCTGT  
ACAGCATGTCTCTTCTCGAAGACAAAGGTCTGGCAGAGGTGTCTAAGGTTGCAGAGCAAGTA  
CTGAATGCCGTAAATAAGGGGCTCTACAGAGAGGCCACAGAGCTGTGGGGGAAAGCAGAAAT  
GATCATTGAACAGAACACAGATGGGGTGAACCTTCTATAACATCTTAACCTAAAAGCACTCCCA  
CGTCTACAATGGAGTCGAGTCTAGAATTACACAGAGCCACCTAGTTTGTCTTTGTCAGCGC  
CACGTGAGACACCTACAACGAGATGCCTTAAGCCAGCTCATGAATGGCCCCATCAGAAAGAA  
GCTCAAAATTATTCCTGAGGATCAATCCTGGGGAGGCCAGGCTACCAACGTCTTTGTGAACA  
TGGAGGAGGACTTCATGAAGCCAGTCATTAGCATTGTGGACGAGTTGCTGGAGGCAGGGATC  
AACGTGACGGTGTATAATGGACAGCTGGATCTCATCGTAGATACCATGGGTCAGGAGGCCTG  
GGTGCGGAAACTGAAGTGGCCAGAACTGCCTAAATTTCAGTCAGCTGAAGTGGAAAGGCCCTGT  
ACAGTGACCCTAAATCTTTGGAAACATCTGCTTTTGTCAAGTCCTACAAGAACCTTGCTTTC  
TACTGGATTCTGAAAGCTGGTCATATGGTTCTTCTGACCAAGGGGACATGGCTCTGAAGAT  
GATGAGACTGGTGAAGTCAAGCAAGAATAGGATGGATGGGGCTGGAGATGAGCTGGTTTGGCCT  
TGGGGCACAGAGCTGAGCTGAGGCCGCTGAAGCTGTAGGAAGCGCCATTCTTCCCTGTATCT  
AACTGGGGCTGTGATCAAGAAGGTCTGACCAGCTTCTGCAGAGGATAAAATCATTGTCTCT  
GGAGGCAATTTGGAAATTATTTCTGCTTCTTAAAAAACCTAAGATTTTAAAAAATTGAT  
TTGTTTTGATCAAAATAAAGGATGATAATAGATATTAA

## **FIGURE 90**

MELALRRSPVPRWLLLLPLLLGLNAGAVIDWPTEEGKEVWDYVTVRKDAYMFWWLYYATNSC  
KNFSELPLVMWLQGGPGGSSTGFGNFEEIGPLDSLKPRKTTWLQAASLLFVDNPVGTGFSY  
VNGSGAYAKDLAMVASDMMVLLKTFFSCHKEFQTVPFYIFSESYGGKMAAGIGLELYKAIQR  
GTIKCNFAGVALGDSWISPVDSVLSWGPYLYSMSLLEDKGLAEVSKVAEQVLNAVKNKGLYRE  
ATELWGKAEMII EQNTDGVNFYNILTKSTPTSTMESSLEFTQSHLVCLCQRHVRHLQORDALS  
QLMNGPIRKKLKII PEDQSWGGQATNVFVNMEEDFMKPVISIVDELLEAGINVTVYNGQLDL  
IVDTMGQEAWVRKLKWPPELPKFSQLKWKALYSDPKSLETSAFVKSYKNLAFYWILKAGHMVP  
SDQGDMAKMMRLVTQQE

### **Signal sequence:**

amino acids 1-25

### **N-glycosylation site.**

amino acids 64-68, 126-130, 362-366

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 101-105

### **Casein kinase II phosphorylation site.**

amino acids 204-208, 220-224, 280-284, 284-288, 351-355, 449-453

### **N-myristoylation site.**

amino acids 22-28, 76-82, 79-85, 80-86, 119-125, 169-175,  
187-193, 195-201, 331-337, 332-338, 360-366

## **FIGURE 91**

GGCCGCGGGAGAGGAGGCCATGGCGCGCGCGGGCGCTGCTGCTGGCGCTGCTGCTGGCTC  
GGGCTGGACTCAGGAAGCCGGAGTCGCAGGAGGCGGCGCCGTTATCAGGACCATGCGGCCGA  
CGGGTCATCACGTCGCGCATCGTGGGTGGAGAGGACGCCGAACCTCGGGCGTTGGCCGTGGCA  
GGGGAGCCTGCGCCTGTGGGATTCCCACGTATGCGGAGTGAGCCTGCTCAGCCACCGCTGGG  
CACTCACGGCGGCGCACTGCTTTGAAACCTATAGTGACCTTAGTGATCCCTCCGGGTGGATG  
GTCCAGTTTGGCCAGCTGACTTCCATGCCATCCTTCTGGAGCCTGCAGGCCTACTACACCCG  
TTACTTCGTATCGAATATCTATCTGAGCCCTCGCTACCTGGGGAATTCACCCTATGACATTG  
CCTTGGTGAAGCTGTCTGCACCTGTCACCTACACTAAACACATCCAGCCCATCTGTCTCCAG  
GCCTCCACATTTGAGTTTGAGAACCGGACAGACTGCTGGGTGACTGGCTGGGGGTACATCAA  
AGAGGATGAGGCACTGCCATCTCCCCACACCCTCCAGGAAGTTCAGGTCGCCATCATAAACA  
ACTCTATGTGCAACCACCTCTTCCTCAAGTACAGTTTCCGCAAGGACATCTTTGGAGACATG  
GTTTGTGCTGGCAACGCCCAAGGCGGGAAGGATGCCTGCTTCGGTGACTCAGGTGGACCCTT  
GGCCTGTAACAAGAATGGACTGTGGTATCAGATTGGAGTCGTGAGCTGGGGAGTGGGCTGTG  
GTCGGCCCAATCGGCCCCGGTGTCTACACCAATATCAGCCACCACTTTGAGTGGATCCAGAAG  
CTGATGGCCCAGAGTGGCATGTCCCAGCCAGACCCCTCCTGGCCACTACTCTTTTTCCCTCT  
TCTCTGGGCTCTCCCACTCCTGGGGCCGGTCTTGAGCCTACCTGAGCCCATGCAGCCTGGGGC  
CACTGCCAAGTCAGGCCCTGGTTCTCTTCTGTCTTGTTTGGTAATAAACACATTCCAGTTGA  
TGCCTTGCAGGGCATTCTTCAAAAAAAAAAAAAAAAAAAAAAAAAA

## **FIGURE 92**

MGARGALLLALLLARAGLRKPESQEAAPLSGPCGRRVITSRIVGGEDAELGRWPWQGSRLRW  
DSHVCGVSLLSHRWALTAAHCFETYSDLSDPSGWMVQFGQLTSMPSFWSLQAYYTRYFVSNI  
YLSPRYLGNSPYDIALVKLSAPVTYTKHIQPICLQASTFEFENRTDCWVTGWGYIKEDEALP  
SPHTLQEVQVAIINNSMCNHLFLKYFRKDI FGDMVCAGNAQGGKDACFGDSGGPLACNKNG  
LWYQIGVVSWGVGCGRPNRPGVYTNISHHFEWIQKLMAQSGMSQPDPSWPLLFFPLLWALPL  
LGPV

### **Signal sequence:**

amino acids 1-18

### **N-glycosylation site.**

amino acids 167-171, 200-204, 273-277

### **Casein kinase II phosphorylation site.**

amino acids 86-90, 134-138, 161-165, 190-194, 291-295

### **N-myristoylation site.**

amino acids 2-8, 44-50, 101-107, 225-231, 229-235, 239-245,  
259-265, 269-275

### **Amidation site.**

amino acids 33-37

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 252-263,

### **Serine proteases, trypsin family, histidine active site.**

amino acids 78-84

biochem



## FIGURE 93

CCCACGCGTCCGCGGACGCGTGGGAAGGGCAGAAATGGGACTCCAAGCCTGCCTCCTAGGGCT  
CTTTGCCCTCATCCTCTCTGGCAAATGCAGTTACAGCCCGAGCCCGACCAGCGGAGGACGC  
TGCCCCCAGGCTGGGTGTCCCTGGGCGGTGCGGACCCTGAGGAAGAGCTGAGTCTCACCTTT  
GCCCTGAGACAGCAGAATGTGGAAAGACTCTCGGAGCTGGTGCAGGCTGTGTGCGGATCCCAG  
CTCTCCTCAATACGGAAAATACCTGACCCTAGAGAATGTGGCTGATCTGGTGAGGCCATCCC  
CACTGACCCTCCACACGGTGCAAAAATGGCTCTTGGCAGCCGGAGCCCAGAAAGTGCCATTCT  
GTGATCACACAGGACTTTCTGACTTGCTGGCTGAGCATCCGACAAGCAGAGCTGCTGCTCCC  
TGGGGCTGAGTTTCATCACTATGTGGGAGGACCTACGAAACCCATGTTGTAAGGTCCCCAC  
ATCCCTACCAGCTTCCACAGGCCTTGCCCCCCCATGTGGACTTTGTGGGGGGACTGCACCGT  
TTTCCCCCAACATCATCCCTGAGGCAACGTCCTGAGCCGCAGGTGACAGGGACTGTAGGCCT  
GCATCTGGGGGTAAACCCCTCTGTGATCCGTAAGCGATACAACTTGACCTCACAAGACGTGG  
GCTCTGGCACCAGCAATAACAGCCAAGCCTGTGCCAGTTTCTGGAGCAGTATTTCCATGAC  
TCAGACCTGGCTCAGTTTATGCGCCTCTTCGGTGGCAACTTTGCACATCAGGCATCAGTAGC  
CCGTGTGGTTGGACAACAGGGCCGGGGCCGGGCGGGGATTGAGGCCAGTCTAGATGTGCAGT  
ACCTGATGAGTGCTGGTGCCAACATCTCCACCTGGGTCTACAGTAGCCCTGGCCGGCATGAG  
GGACAGGAGCCCTTCTGCAGTGGCTCATGCTGCTCAGTAATGAGTCAGCCCTGCCACATGT  
GCATACTGTGAGCTATGGAGATGATGAGGACTCCCTCAGCAGCGCCTACATCCAGCGGGTCA  
ACACTGAGCTCATGAAGGCTGCCGCTCGGGGTCTCACCTGCTCTTCGCCTCAGGTGACAGT  
GGGGCCGGGTGTTGGTCTGTCTCTGGAAGACACCAGTTCCGCCCTACCTTCCCTGCCTCCAG  
CCCCATGTCAACACAGTGGGAGGCACATCCTTCCAGGAACCTTTCTCATCACAATGAAA  
TTGTTGACTATATCAGTGGTGGTGGCTTCAGCAATGTGTTCCACGGCCTTCATACCAGGAG  
GAAGCTGTAACGAAGTTCTGAGCTCTAGCCCCACCTGCCACCATCCAGTTACTTCAATGC  
CAGTGGCCGTGCCTACCCAGATGTGGCTGCACCTTTCTGATGGCTACTGGGTGGTCAACA  
GAGTGCCCATTTCCATGGGTGTCCGGAACCTCGGCCTCTACTCCAGTGTTTGGGGGGATCCTA  
TCCTTGATCAATGAGCACAGGATCCTTAGTGCCCGCCCCCTCTTGGCTTTCTCAACCCAAG  
GCTCTACCAGCAGCATGGGGCAGGTCTCTTTGATGTAACCCGTGGCTGCCATGAGTCCTGTC  
TGGATGAAGAGGTAGAGGGCCAGGGTTTCTGCTCTGGTCTGGCTGGGATCCTGTAACAGGC  
TGGGGAACACCAACTTCCCAGCTTTGCTGAAGACTCTACTCAACCCCTGACCCTTTCTATC  
AGGAGAGATGGCTTGTCCCTGCCCTGAAGCTGGCAGTTTCTAGTCCCTTATTCTGCCCTGTTG  
GAAGCCCTGCTGAACCCCTCAACTATTGACTGCTGCAGACAGCTTATCTCCCTAACCCTGAAA  
TGCTGTGAGCTTGACTTGACTCCCAACCCTACCATGCTCCATCATACTCAGGTCTCCCTACT  
CCTGCCCTTAGATTCTCAATAAGATGCTGTAAGTAGCATTTTTTTGAATGCCTCTCCCTCCGC  
ATCTCATCTTTCTCTTTTCAATCAGGCTTTTCCAAAGGGTTGTATACAGACTCTGTGCACTA  
TTTCACTTGATATTCAATCCCCAATTCAGTGAAGGAGACCTCTACTGTCACCGTTTACTCT  
TTCCTACCCTGACATCCAGAAACAATGGCCTCCAGTGCATACTTCTCAATCTTTGCTTTATG  
GCCTTTCCATCATAGTTGCCCACTCCCTCTCCTTACTTAGCTTCCAGGTCTTAACTTCTCTG  
ACTACTCTTGTCTTCTCTCATCAATTTCTGCTTCTTCATGGAATGCTGACCTTCATTGC  
TCCATTTGTAGATTTTTGCTCTTCTCAGTTTACTCATTGTCCCCTGGAACAAATCACTGACA  
TCTACAACCATTACCATCTCACTAAATAAGACTTTCTATCCAATAATGATTGATACCTCAAA  
TGTAACAAA

## **FIGURE 94**

MGLQACLLGLFALILSGKCSYSPEPDQRRTLPPGWVSLGRADPEEELSLTFALRQQNVERLS  
ELVQAVSDPSSPQYGKYLTLENVADLVRPSPLTLHTVQKWLLAAGAQKCHSVITQDFLTCWL  
SIRQAELLPLGAEFHHYVGGPTETHVVRSPHPYQLPQALAPHVDFVGGLHRFPPTSSLRQRP  
EPQVTGTVGLHLGVTPSVIRKRYNLTSQDVGSGTSNNSQACAQFLEQYFHDSDLAQFMRLFG  
GNFAHQASVARVVGQQGRGRAGIEASLDVQYLMSAGANISTWVYSSPGRHEGQEPFLQWLML  
LSNESALPHVHTVSYGDDEDSLSSAYIQRVNTELMKAAARGLTLLFASGDSGAGCWSVSGRH  
QFRPTFPASSPYVTTVGGTSFQEPFLITNEIVDYISGGGFSNVFPRPSYQEEAVTKFLSSSP  
HLPPSSYFNASGRAYPDVAALSDGYWVVSNRVPIPWVSGTSASTPVFGGILSLINEHRILSG  
RPPLGLNPRLYQQHGAGLFDVTRGCHESCLDEEVEGQGFCSGPGWDPVTGWGTPTSQLC

### **Signal sequence:**

amino acids 1-16

### **N-glycosylation site.**

amino acids 210-214, 222-226, 286-290, 313-317, 443-447

### **Glycosaminoglycan attachment site.**

amino acids 361-365, 408-412, 538-542

### **Casein kinase II phosphorylation site.**

amino acids 212-216, 324-328, 392-396, 420-424, 525-529

### **N-myristoylation site.**

amino acids 2-8, 107-113, 195-201, 199-205, 217-223, 219-225,  
248-254, 270-276, 284-290, 409-415, 410-416, 473-479, 482-488,  
521-527, 533-539, 549-555



## **FIGURE 96**

MAGIPGLLFLFFLLCAVGQVSPYSAPWKPTWPAYRLPVVLPQSTLNLAKPDFGAEAKLEVS  
SSCGPQCHKGTPLPTYEEAKQYLSYETLYANGSRTETQVGIYILSSSGDGAQHRDSGSSGKS  
RRKRQIYGYSRFSIFGKDFLLNYPFSTSVKLSTGCTGTLVAEKHVLTAAHCIHDGKTYVKG  
TQKLRVGFLKPKFKDGGRGANDSTSAMPEQMKFQWIRVKRTHVPKGWIKGNANDIGMDYDYA  
LLELKKPHKRKFMKIGVSPPAKQLPGGRIHFSGYDNDRPGNLVYRFCDVKDETYDLLYQQCD  
AQPASGSGVYVRMWKRQQQKWERKIIGIFSGHQWVDMNGSPQDFNVAVRITPLKYAQICYW  
IKGNYLDCREG

### **Signal sequence:**

amino acids 1-19

### **N-glycosylation site.**

amino acids 93-97, 207-211

### **Glycosaminoglycan attachment site.**

amino acids 109-113, 316-320

### **Casein kinase II phosphorylation site.**

amino acids 77-81, 95-99, 108-112, 280-284, 351-355

### **N-myristoylation site.**

amino acids 159-165, 162-168, 202-208, 205-211, 314-320, 338-344

### **Serine proteases, trypsin family, histidine active site.**

amino acids 171-177

protein database

## **FIGURE 97**

GCATCGCCCTGGGTCTCTCGAGCCTGCTGCCTGCTCCCCGCCCCACCAGCCATGGTGGTTT  
CTGGAGCGCCCCAGCCCTGGGTGGGGGCTGTCTCGGCACCTTCACCTCCCTGCTGCTGCTG  
GCGTCGACAGCCATCCTCAATGCGGCCAGGATACCTGTTCCCCCAGCCTGTGGGAAGCCCCA  
GCAGCTGAACCGGGTTGTGGGCGGCGAGGACAGCACTGACAGCGAGTGGCCCTGGATCGTGA  
GCATCCAGAAGAATGGGACCCACCACTGCGCAGGTTCTCTGCTCACCAGCCGCTGGGTGATC  
ACTGCTGCCCCACTGTTTTCAAGGACAACCTGAACAAACCATACTGTTCTCTGTGCTGCTGGG  
GGCCTGGCAGCTGGGGAACCCTGGCTCTCGGTCCCAGAAGGTGGGTGTTGCCTGGGTGGAGC  
CCCACCCTGTGTATTCTCTGGAAGGAAGGTGCCTGTGCAGACATTGCCCTGGTGCCTCTCGAG  
CGCTCCATACAGTTCTCAGAGCGGGTCCTGCCCATCTGCCTACCTGATGCCTCTATCCACCT  
CCCTCCAAACACCCACTGCTGGATCTCAGGCTGGGGGAGCATCCAAGATGGAGTTCCCTTGC  
CCCACCCTCAGACCCTGCAGAAGCTGAAGGTTCTATCATCGACTCGGAAGTCTGCAGCCAT  
CTGTACTGGCGGGGAGCAGGACAGGGACCCATCACTGAGGACATGCTGTGTGCCGGCTACTT  
GGAGGGGGAGCGGGATGCTTGTCTGGGCGACTCCGGGGGCCCCCTCATGTGCCAGGTGGACG  
GCGCCTGGCTGCTGGCCGGCATCATCAGCTGGGGCGAGGGCTGTGCCGAGCGCAACAGGCCC  
GGGGTCTACATCAGCCTCTCTGCGCACCGCTCCTGGGTGGAGAAGATCGTGCAAGGGGTGCA  
GCTCCGCGGGCGCGCTCAGGGGGGTGGGGCCCTCAGGGCACCGAGCCAGGGCTCTGGGGCCG  
CCGCGCGCTCCTAGGGCGCAGCGGGACGCGGGGCTCGGATCTGAAAGGCGGCCAGATCCACA  
TCTGGATCTGGATCTGCGGCGGCCTCGGGCGGTTTCCCCCGCCGTAAATAGGCTCATCTACC  
TCTACCTCTGGGGGCCCCGACGGCTGCTGCGGAAAGGAAACCCCTCCCCGACCCGCCCCGAC  
GGCCTCAGGCCCCCTCCAAGGCATCAGGCCCCGCCAACGGCCTCATGTCCCCGCCCCCAC  
GACTTCCGGCCCCGCCCCGGGCCCCAGCGCTTTTGTGTATATAAATGTTAATGATTTTTAT  
AGGTATTTGTAACCCTGCCACATATCTTATTTATTCTCCAATTTCAATAAATTATTTATT  
CTCCAAAAAAAAAA

bioRxiv preprint doi: <https://doi.org/10.1101/000000>; this version posted January 1, 2014. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

## **FIGURE 98**

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA43318

><subunit 1 of 1, 317 aa, 1 stop

><MW: 33732, pI: 7.90, NX(S/T): 1

MVMSGAPPALGGGCLGTFTSLLLLASTAILNAARIPVPPACGKPQQLNRVVGGEDSTDSEWP  
WIVSIQKNGTHHCAGSLLTSRWVITAAHCFKDNLNKPYLFSVLLGAWQLGNPGSRSQKVGVA  
WVEPHPVYSWKEGACADIALVRLERSIQFSERVLPICLPDASIHLPNTHCWISGWGSIQDG  
VPLPHPQTLQKLKVPIIDSEVCSHLYWRGAGQGPITEDMLCAGYLEGERDACLGDSGGPLMC  
QVDGAWLLAGIISWGEGCAERNRPGVYISLSAHRSWVEKIVQGVQLRGRAQGGGALRAPSQG  
SGAAARS

### **Signal sequence:**

amino acids 1-32

### **N-glycosylation site.**

amino acids 62-66, 96-100, 214-218, 382-386, 409-413, 455-459,  
628-632, 669-673, 845-849, 927-931, 939-943, 956-960

### **Glycosaminoglycan attachment site.**

amino acids 826-830

### **Casein kinase II phosphorylation site.**

amino acids 17-21, 39-43, 120-124, 203-207, 254-258, 264-268,  
314-318, 323-327, 347-351, 464-468, 548-552, 632-636, 649-653,  
671-675, 739-743, 783-787, 803-807, 847-851, 943-947, 958-962,  
1013-1017, 1019-1023, 1021-1025

### **Tyrosine kinase phosphorylation site.**

amino acids 607-615

### **N-myristoylation site.**

amino acids 179-185, 197-203, 320-326, 367-373, 453-459, 528-534,  
612-618, 623-629, 714-720, 873-879

The first two columns of the table are the names of the authors and the year of publication. The third column is the title of the paper. The fourth column is the journal name. The fifth column is the volume number. The sixth column is the page number. The seventh column is the year of publication.

GACGGCTGGCCACCAATGCACGGCTCCTGCAGTTTCCTGATGCTTCTGCTGCCGCTACTGCTA  
CTGCTGGTGGCCACCACAGGCCCGTTGGAGCCCTCACAGATGAGGAGAAACGTTTGTATGGT  
GGAGCTGCACAACCTCTACCGGGCCCAGGTATCCCCGACGGCCTCAGACATGCTGCACATGA  
GATGGGACGAGGAGCTGGCCGCCTTCGCCAAGGCCTACGCACGGCAGTGCCTGTGGGGCCAC  
AACAAGGAGCGCGGGCGCCGCGGCGAGAATCTGTTGCCATCACAGACGAGGGCATGGACGT  
GCCGCTGGCCATGGAGGAGTGGCACCACGAGCGTGAGCACTACAACCTCAGCGCCGCCACCT  
GCAGCCCAGGCCAGATGTGCGGCCACTACACGCAGGTGGTATGGGCCAAGACAGAGAGGATC  
GGCTGTGGTTCCCACTTCTGTGAGAAGCTCCAGGGTGTTGAGGAGACCAACATCGAATTACT  
GGTGTGCAACTATGAGCCTCCGGGGAACGTGAAGGGGAAACGGCCCTACCAGGAGGGGACTC  
CGTGCTCCCAATGTCCCTCTGGCTACCACTGCAAGAACTCCCTCTGTGAACCCATCGGAAGC  
CCGGAAGATGCTCAGGATTTGCCTTACCTGGTAACTGAGGCCCCATCCTTCCGGGCGACTGA  
AGCATCAGACTCTAGGAAAATGGGTACTCCTTCTTCCCTAGCAACGGGGATTCCGGCTTTCT  
TGGTAACAGAGGTCTCAGGCTCCCTGGCAACCAAGGCTCTGCCTGCTGTGGAACCCAGGCC  
CCAACTTCCTTAGCAACGAAAGACCCGCCCTCCATGGCAACAGAGGCTCCACCTTGCGTAAC  
AACTGAGGTCCCTTCCATTTTGGCAGCTCACAGCCTGCCCTCCTTGGATGAGGAGCCAGTTA  
CCTTCCCCAAATCGACCCATGTTCCCTATCCCAAATCAGCAGACAAAGTGACAGACAAAACA  
AAAGTGCCCTCTAGGAGCCCAGAGAACTCTCTGGACCCCAAGATGTCCCTGACAGGGGCAAG  
GGAACTCCTACCCCATGCCCAGGAGGAGGCTGAGGCTGAGGCTGAGTTGCCTCCTTCCAGTG  
AGGTCTTGGCCTCAGTTTTTCCAGCCCAGGACAAGCCAGGTGAGCTGCAGGCCACACTGGAC  
CACACGGGGCACACCTCCTCCAAGTCCCTGCCCAATTTCCCCAATACCTCTGCCACCGCTAA  
TGCCACGGGTGGGCGTGCCCTGGCTCTGCAGTCGTCTTGCCAGGTGCAGAGGGCCCTGACA  
AGCCTAGCGTTGTGTGTCAGGGCTGAACTCGGGCCCTGGTCATGTGTGGGGCCCTCTCCTGGGA  
CTACTGCTCCTGCCTCCTCTGGTGTTGGCTGGAATCTTCTGAATGGGATACCACTCAAAGGG  
TGAAGAGGTCAGCTGTCTCCTGTCTATCTTCCCCACCCTGTCCCCAGCCCCCTAAACAAGATA  
CTTCTTGGTTAAGGCCCTCCGGAAGGGAAAGGCTACGGGGCATGTGCCTCATCACACCATCC  
ATCCTGGAGGCACAAGGCCTGGCTGGCTGCGAGCTCAGGAGGCCGCTGAGGACTGCACACC  
GGGCCCACACCTCTCCTGCCCCCTCCCTCCTGAGTCCTGGGGGTGGGAGGATTTGAGGGAGCT  
CACTGCCTACCTGGCCTGGGGCTGTCTGCCACACAGCATGTGCGCTCTCCCTGAGTGCCTG  
TGTAGCTGGGGATGGGGATTCTAGGGGCAGATGAAGGACAAGCCCCACTGGAGTGGGGTTCT  
TTTGAGTGGGGGAGGCAGGGACGAGGGAAGGAAAGTAACTCCTGACTCTCCAATAAAAACCT  
GTCCAACCTGTGAAA

## **FIGURE 100**

MHGSCSFLMLLLPLLLLLLVATTGPVGALTDEEKRLMVELHNLRYAQVSPTASDMLHMRWDEE  
LAFAKAYARQCVWGHNKERGRGENLFAITDEGMDVPLAMEEWHHEREHYNLSAATCSPGQ  
MCGHYTQVVWAKTERIGCGSHFCEKLGVEETNIELLVLCNYEPPGNVKGKRPYQEGTPCSQC  
PSGYHCKNSLCEPIGSPEDAQDLPYLVTEAPSFRAEASDSRKMGTTPSSLATGIPAFVLVTEV  
SGSLATKALPAVETQAPTSLATKDPPSMATEAPPCVTTEVPSILAAHSLPSLDEEPVTFPKS  
THVPIPKSADKVTDKTKVPSRSPENSLDPKMSLTGARELLPHAQEEAEAEELPPSSEVLAS  
VFPAQDKPGELQATLDHTGHTSSKSLPNFPNTSATANATGGRALALQSSLPGAEGPDKPSVV  
SGLNSGPGHVWGPLLGLLLLLPPLVLAGIF

### **Signal sequence:**

amino acids 1-22

### **N-glycosylation site.**

amino acids 114-118, 403-407, 409-413

### **Glycosaminoglycan attachment site.**

amino acids 439-443

### **Casein kinase II phosphorylation site.**

amino acids 29-33, 50-54, 156-160, 195-199, 202-206, 299-303

### **N-myristoylation site.**

amino acids 123-129, 143-149, 152-158, 169-175, 180-186, 231-237,  
250-256

### **Amidation site.**

amino acids 82-86, 172-176

### **Peroxidases proximal heme-ligand signature.**

amino acids 287-298

### **Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 1.**

amino acids 127-138

### **Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 signature 2.**

amino acids 160-172



# FIGURE 101

GTAAGTGAAGTCAGGCTTTTTCATTTGGGAAGCCCCCTCAACAGAATTTCGGTCATTCTCCAAGTTATGGTGGACGT  
 ACTTCTGTTGTTCTCCCTCTGCTTGTCTTTTTCACATTAGCAGACCGGACTTAAGTCACAACAGATTATCTTTTCAT  
 CAAGGCAAGTTCCATGAGCCACCTTCAAAGCCTTCGAGAAGTGAACTGAACAACAATGAATTGGAGACCATTCC  
 AAATCTGGGACCAGTCTCGGCAAATATTACACTTCTCTCCTTGGCTGGAAACAGGATTGTTGAAATACTCCCTGA  
 ACATCTGAAAGAGTTTCAGTCCCTTGAACTTTGGACCTTAGCAGCAACAATATTTTCAGAGCTCCAACTGCATT  
 TCCAGCCCTACAGCTCAAATATCTGTATCTCAACAGCAACCGAGTCACATCAATGGAACCTGGGTATTTTGACAA  
 TTTGGCCAACACACTCCTTGTGTTAAAGCTGAACAGGAACCGAATCTCAGCTATCCCACCCAAGATGTTTAACT  
 GCCCCAACTGCAACATCTCGAATTGAACCGAAACAAGATTAAAAATGTAGATGGACTGACATTCCAAGGCCTTGG  
 TGCTCTGAAGTCTCTGAAAATGCAAAGAAATGGAGTAACGAAACTTATGGATGGAGCTTTTGGGGGCTGAGCAA  
 CATGGAAATTTTGCAGCTGGACCATAACAACCTAACAGAGATTACCAAAGGCTGGCTTTACGGCTTGCTGATGCT  
 GCAGGAACCTCATCTCAGCCAAAATGCCATCAACAGGATCAGCCCTGATGCCTGGGAGTTCTGCCAGAAGCTCAG  
 TGAGCTGGACCTAACTTTCAATCACTTATCAAGGTTAGATGATTCAAGCTTCTTGGCCTAAGCTTACTAAATAC  
 ACTGCACATTGGGAACAACAGAGTCAGCTACATTGCTGATTGTGCCTTCCGGGGGCTTTCCAGTTTAAAGACTTT  
 GGATCTGAAGAACAATGAAATTTCTGGACTATTGAAGACATGAATGGTGCTTTCTCTGGGCTTGACAACTGAG  
 GCGACTGATACTCCAAGGAAATCGGATCCGTTCTATTACTAAAAAGCCTTCACTGGTTGGATGCATTGGAGCA  
 TCTAGACCTGAGTGACAACGCAATCATGTCTTTACAAGGCAATGCATTTTCACAAATGAAGAACTGCAACAATT  
 GCATTTAAATACATCAAGCCTTTTGTGCGATTGCCAGCTAAAATGGCTCCACAGTGGGTGGCGGAAACAACCTT  
 TCAGAGCTTTGTAAATGCCAGTTGTGCCCATCCTCAGCTGCTAAAAGGAAGAAGCATTTTGTCTGTTAGCCCAGA  
 TGGCTTTGTGTGTGATGATTTTCCCAAACCCAGATCACGGTTTCAGCCAGAAACACAGTCGGCAATAAAGGTTT  
 CAATTTGAGTTTTCATCTGCTCAGCTGCCAGCAGCAGTGATTCCCAATGACTTTTGTCTGGAAAAAGACAATGA  
 ACTACTGCATGATGCTGAAATGGAATATATGCACACCTCCGGGCCCAAGGTGGCGAGGTGATGGAGTATACCAC  
 CATCCTTCGGCTGCGCGAGGTGGAATTTGCCAGTGAGGGGAAATATCAGTGTGTCTCTCCAATCACTTTGGTTT  
 ATCCTACTCTGTCAAAGCCAAGCTTACAGTAAATATGCTTCCCTCATTACCAAGACCCCCATGGATCTCACCAT  
 CCGAGCTGGGGCCATGGCAGCTTGGAGTGTGCTGTCTGTGGGGCACCCAGCCCCCAGATAGCCTGGCAGAAGGA  
 TGGGGGCACAGACTTCCAGCTGCACGGGAGAGACGCGATGCTGTGATGCCCGAGGATGACGTGTTCTTTATCGT  
 GGATGTGAAGATAGAGGACATTGGGGTATACAGCTGCACAGCTCAGAACAGTGCAGGAAGTATTTTCAGCAAATGC  
 AACTCTGACTGTCTTAGAAACACCATCATTTTTGCGGCCACTGTTGGACCGAAGTGAACCAAGGGAGAAACAGC  
 CGTCTTACAGTGCAATTGCTGGAGGAAGCCCTCCCCCTAACTGAACTGGACCAAAGATGATAGCCCATTTGGTGGT  
 AACCAGAGAGGCATTTTTTTCAGCAGGCAATCAGCTTCTGATTATTGTGGACTCAGATGTCAGTGATGCTGGGAA  
 ATACACATGTGAGATGTCTAACACCCCTTGGCACTGAGAGAGGAACGTCGCGCTCAGTGTGATCCCCACTCCAAC  
 CTGCGACTCCCCCTCAGATGACAGCCCCATCGTTAGACGATGACGGATGGGCCACTGTGGGTGTCGTGATCATAGC  
 CGTGGTTTGTGTGTGGTGGGCACGTCACTCGTGTGGGTGGTTCATCATATACCACACAAGGCGGAGGAATGAAGA  
 TTGCAGCATTACCAACACAGATGAGACCAACTTGCCAGCAGATATTCCTAGTTATTTGTCTCAGGGAAACGTT  
 AGCTGACAGGCAGGATGGGTACGTGTCTTCAAGAAAGTGGAAGCCACCACAGTTTGTCTCATCTTCAGGTGCTGG  
 ATTTTTCTTACCACAACATGACAGTAGTGGGACCTGCCATATTGACAATAGCAGTGAAGCTGATGTGGAAGCTGC  
 CACAGATCTGTTCCCTTTGTCCGTTTTTGGGATCCACAGGCCCTATGTATTTGAAGGGAAATGTGTATGGCTCAGA  
 TCCTTTTGAACATATCATACAGGTTGCAGTCTGACCCAAGAACAGTTTAAATGGACCACTATGAGCCCAGTTA  
 CATAAAGAAAAAGGAGTGCTACCCATGTTCTCATCTTCAAGAAATCCTGCGAACGGAGCTTCAGTAATATATC  
 GTGGCCTTTCATATGTGAGGAAGCTACTTAACACTAGTTACTCTCAATGAAGGACCTGGAATGAAAAATCTGTG  
 TCTAAACAAGTCCCTCTTAGATTTTAGTGCAAATCCAGAGCCAGCGTCGGTTGCCTCGAGTAATCTTTTCATGGG  
 TACCTTTGGAAAAGCTCTCAGGAGACCTCACCTAGATGCCTATTCAAGCTTTGGACAGCCATCAGATTGTCAGCC  
 AAGAGCCTTTTATTTGAAAGCTCATTCTTCCCAGACTTGGACTCTGGGTGAGAGGAAGATGGGAAAGAAAGGAC  
 AGATTTTCAGGAAGAAAATCACATTTGTACCTTTAAACAGACTTTAGAAAACACAGGACTCCAAATTTTCAGTC  
 TTATGACTTTGGACACATAGACTGAATGAGACCAAAGGAAAAGCTTAACATACTACCTCAAGTGAACCTTTTATTTA  
 AAAGAGAGAGAATCTTATGTTTTTAAATGGAGTTATGAATTTTAAAGGATAAAAATGCTTTATTTATACAGAT  
 GAACCAAAATTAACAAAAGTTATGAAAATTTTATACTGGGAATGATGCTCATATAAGAATACCTTTTAAACTA  
 TTTTAACTTTGTTTTATGCAAAAAGTATCTTACGTAAATTAATGATATAAATCATGATTATTTATGTATTT  
 TTATAATGCCAGATTTCTTTTTATGGAAAATGAGTTACTAAAGCATTTTAAATAATACCTGCTTGTACCATTTT  
 TTAATAAGAAGTTACTTCATTATATTTTGCACATTATATTTAATAAATGTGTCAATTTGAA

Sequence 101

## **FIGURE 102**

MVDVLLLFSLCLLFHISRPDLSHNRLSFIKASSMSHLQSLREVKLNNNELETIPNLGPVSAN  
ITLLSLAGNRIVEILPEHLKEFQSLETLDLSSNNISELQTAFPALQLKYLYLNSNRVTSMEP  
GYFDNLANTLLVLKLNRRISAIPPKMFKLPLQHLLELNRNKIKNVDGLTFQGLGALKSLKM  
QRNGVTKLMDGAFWGLSNMEILQLDHNNLTEITKGWLYGLLMLQELHLSQNAINRISPDAWE  
FCQKLSELDLTFNHL SRLDDSSFLGLSLNLTLHIGNNRVSYIADCAFRGLSSLKTLDLKNNE  
ISWTIEDMNGAFSGLDKLRLILQGNRIRSITKKAFTGLDALEHLDLSDNAIMSLQGNAFSQ  
MKKLQQLHLNTSSLLCDCQLKWLPQWVAENNFSQSFVNASCAHPQLLKGRSIFAVSPDGFVCD  
DFPKPQITVQPETQSAIKGSNLSFICSAASSSDSPMTFAWKDNELLHDAEMENY AHLRAQG  
GEVMEYTTILRLREVEFASEGKYQCVISNHFGSSYSVKAKLTVNMLPSFTKTPMDLTIRAGA  
MARLECAAVGHPAPQIAWQKGGTDFPAARERRMHVMPEDDVFFIVDVKIEDIGVYSCTAQN  
SAGSISANATLTVLETSPFLRPLLDRTVTKGETAVLQCIAGGSPPPKNWTKDDSPLVVTER  
HFFAAGNQLLIIVDSVDAGKYTCEMSNTLGTGERGNVRLSVIPTPTCDSPQMTAPSLDDDG  
WATVGVVIIAVVCCVVGTSLVWVVIYHTRRRNEDCSITNTDETNPADIPSYLSSQGTLD  
RQDGYVSSESGSHHQFVTSSGAGFFLPQHDSSGTCHIDNSSEADVEAATDLFLCPFLGSTGP  
MYLKGNVYGSDPFETYHTGCSPPDRTVLMDHYEPSYIKKKECYPCSHPSEESCERSFSNISW  
PSHVRKLLNTSYSHNEGPGMKNLCLNKSSLD FSANPEPASVASSNSFMGTFGKALRRPHLDA  
YSSFGQPSDCQPRAFYLKAHSSPDLDGSEEDGKERTDFQEEHNICTFKQTLNRYRTPNFQS  
YDLDT

**Signal sequence:**

amino acids 1-19

**Transmembrane domain:**

amino acids 746-765

**N-glycosylation site.**

amino acids 62-66, 96-100, 214-220, 382-386, 409-413, 455-459,  
628-632, 669-673, 845-849, 927-931, 939-943, 956-960

**Glycosaminoglycan attachment site.**

amino acids 826-830

**Casein kinase II phosphorylation site.**

amino acids 17-21, 39-43, 120-124, 203-207, 254-258, 264-268,  
314-318, 323-327, 347-351, 464-468, 548-552, 632-636, 649-653,  
671-675, 739-743, 783-787, 803-807, 847-851, 943-947, 958-962,  
1013-1017, 1019-1023, 1021-1025

**Tyrosine kinase phosphorylation site.**

amino acids 607-615

**N-myristoylation site.**

amino acids 179-185, 197-203, 320-326, 367-373, 453-459, 528-534,  
612-618, 623-629, 714-720, 873-879

## FIGURE 103

GGGGAGAGGAATTGACCATGTAAAAGGAGACTTTTTTTTTTTGGTGGTGGTGGCTGTTGGGTGCCTTGCAAAAATG  
AAGGATGCAGGACGCAGCTTTCTCCTGGAACCGAACGCAATGGATAAACTGATTGTGCAAGAGAGAAGGAAGAAC  
GAAGCTTTTTCTTGTGAGCCCTGGATCTTAACACAAATGTGTATATGTGCACACAGGGAGCATTCAGAATGAAA  
TAAACCAGAGTTAGACCCGCGGGGGTGGTGTGTTCTGACATAAATAAATAATCTTAAAGCAGCTGTTCCCTCC  
CCACCCCCAAAAAAGGATGATTGGAAATGAAGAACCGAGGATTCACAAAGAAAAAGTATGTTTCATTTTTCTC  
TATAAAGGAGAAAGTGAGCCAAGGAGATATTTTTGGAATGAAAAGTTTGGGGCTTTTTTAGTAAAGTAAAGAACT  
GGTGTGGTGGTGTTCCTTTCTTTTGAATTTCCACAAAGAGGAGAGGAAATTAATAATACATCTGCAAAGAAA  
TTTCAGAGAAGAAAAGTTGACCGCGCAGATTGAGGCATTGATTGGGGGAGAGAAAACCAGCAGAGCACAGTTGGA  
TTTGTGCCATGTTGACTAAAATTGACGGATAATTGCAGTTGGATTTTTCTTCATCAACCTCCTTTTTTTTAAAT  
TTTTATTCTTTTGGTATCAAGATCATGCGTTTTCTCTTGTCTTAACACCTGGATTTCCATCTGGATGTTGCT  
GTGATCAGTCTGAAATACAACCTGTTTGAATTCAGAAGGACCAACACCAGATAAATTATGAATGTTGAACAAGAT  
GACCTTACATCCACAGCAGATAATGATAGGTCCTAGGTTTAAACAGGGCCCTATTTGACCCCTGCTTGTGGTGCT  
GCTGGCTCTTCAACTTCTTGTGGTGGCTGGTCTGGTGCAGGCTCAGACCTGCCCTTCTGTGTGCTCCTGCAGCAA  
CCAGTTCAGCAAGGTGATTTGTGTTCCGAAAAACCTGCGTGAGGTTCCGGATGGCATCTCCACCAACACACGGCT  
GCTGAACCTCCATGAGAACCATAATCCAGATCATCAAAGTGAACAGCTTCAAGCACTTGAGGCACTTGGAATCCT  
ACAGTTGAGTAGGAACCATATCAGAACCATTGAAATTGGGGCTTTCAATGGTCTGGCGAACCTCAACACTCTGGA  
ACTCTTTGACAATCGTCTTACTACCATCCCGAATGGAGCTTTTGTATACTTGTCTAACTGAAGGAGCTCTGGTT  
GCGAAACAACCCATTGAAAGCATCCCTTCTTATGCTTTTAAACAGAATTCCTTCTTTGCGCCGACTAGACTTAGG  
GGAATTGAAAAGACTTTCATACATCTCAGAAGGTGCCTTTGAAGGTCTGTCCAACCTTGAGGTATTTGAACCTTGC  
CATGTGCAACCTTCGGGAAATCCCTAACCTCACACCGCTCATAAACTAGATGAGCTGGATCTTTCTGGGAATCA  
TTTATCTGCCATCAGGCCTGGCTCTTTCCAGGGTTTGATGCACCTTCAAAAACCTGTGGATGATACAGTCCAGAT  
TCAAGTGATTGAACGGAATGCCTTTGACAACCTTCAGTCACTAGTGGAGATCAACCTGGCACACAATAATCTAAC  
ATTACTGCCCTCATGACCTCTTCACTCCCTTGCCATCATCTAGAGCGGATACATTTACATCACAACCCTTGGAACCTG  
TAACTGTGACATACTGTGGCTCAGCTGGTGGATAAAAGACATGGCCCCCTCGAACACAGCTTGTGTGCCCGGTG  
TAACACTCCTCCCAATCTAAAGGGGAGGTACATTGGAGAGCTCGACCAGAATTACTTCACATGCTATGCTCCGGT  
GATTGTGGAGCCCCCTGCAGACCTCAATGTCACTGAAGGCATGGCAGCTGAGCTGAAATGTCGGGCCTCCACATC  
CCTGACATCTGTATCTTGGATTACTCCAAATGGAACAGTCATGACACATGGGGCGTACAAAGTGCAGGATAGCTGT  
GCTCAGTGATGGTACGTTAAATTTACAAATGTAACCTGTGCAAGATACAGGCATGTACACATGTATGGTGAGTAA  
TTCCGTTGGGAATACTACTGCTTCAGCCACCCTGAATGTTACTGCAGCAACCACTACTCCTTCTCTTACTTTTC  
AACCGTCACAGTAGAGACTATGGAACCGTCTCAGGATGAGGCACGGACCAAGATAACAATGTGGGTCCCCTCC  
AGTGGTGCAGTGGGAGACCACCAATGTGACCACCTCTCTCACACCACAGAGCACAAGGTGACAGAGAAAACCTT  
CACCATCCCAGTGACTGATATAAACAGTGGGATCCCAGGAATTGATGAGGTGATGAAGACTACCAAATCATCAT  
TGGGTGTTTTGTGGCCATCACACTCATGGCTGCAGTGATGCTGGTCATTTTCTACAAGATGAGGAAGCAGACCA  
TCGGCAAAACCATCACGCCCCAACAAGGACTGTTGAAATTATTAATGTGGATGATGAGATTACGGGAGACACACC  
CATGGAAAGCCACCTGCCATGCCTGCTATCGAGCATGAGCACCTAAATCACTATAACTCATACAAATCTCCCTT  
CAACCACACAACAACAGTTAAACACAATAAATTCAATACACAGTTTCAGTGCATGAACCGTTATTGATCCGAATGAA  
CTCTAAAGACAATGTACAAGAGACTCAAATCTAAACATTTACAGAGTTACAAAAACAAACAATCAAAAAAAA  
GACAGTTTATTAAAAATGACACAAATGACTGGGCTAAATCTACTGTTTCAAAAAGTGTCTTTACAAAAAACAA  
AAAAGAAAAGAAATTTATTTATTAAAAATTCTATTGTGATCTAAAGCAGACAAAAA

## **FIGURE 104**

MLNKMTLHPQQIMIGPRFNRALFDPLLVLALLQLLVVAGLVRAQTCPSVCSCSNQFSKVIC  
VRKNLREVPDGI STNTRLLNLHENQIQIIKVNSFKHLRHLEILQLSRNHIRTIEIGAFNGLA  
NLNTLELFDNRLTTIPNGAFVYLSKLELWLRNPNPIESIPSYAFNRIPSLRRLDLGELKRLS  
YISEGAFEGLSNLRYLNLAMCNLREIPNLTPLIKLELDL SGNHLSAIRPGSFQGLMHLQKL  
WMIQSQIQVIERNAFDNLQSLVEINLAHNNLTLLPHDLFTPLHHLERIHLLHNPWNCNCDIL  
WLSWWIKDMAPSNTACCARCNTPPNLKGRYIGELDQNYFTCYAPVIVEPPADLNVTEGMAAE  
LKCRASLTSLTSVSWITPNGTVMTHGAYKVRIAVLSDGTLNFTNVTQDTGMYTCMVSNVSGN  
TTASATLNVTAATTTTPFSYFSTVTVETMEPSQDEARTTDNNVGPTPVVDWETTNVTTS LTPQ  
STRSTKFTTIPVTDINSGIPGIDEVMKTTKIIIGCFVAITLMAAVMLVIFYKMRKQHHRQN  
HHAPTRTVEIINVDDEITGDTPMESHLPMPAIEHEHLNHYSYKSPFNHTTTVNTINSIHSS  
VHEPLLIRMNSKDNVQETQI

### **Signal sequence:**

amino acids 1-44

### **Transmembrane domain:**

amino acids 523-543

### **N-glycosylation site.**

amino acids 278-282, 364-368, 390-394, 412-416, 415-419, 434-438,  
442-446, 488-492, 606-610

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 183-187

### **Casein kinase II phosphorylation site.**

amino acids 268-272, 417-421, 465-469, 579-583, 620-624

### **N-myristoylation site.**

amino acids 40-46, 73-79, 118-124, 191-197, 228-234, 237-243,  
391-397, 422-428, 433-439, 531-537

“Biochem”

# FIGURE 105

AGCCGACGCTGCTCAAGCTGCAACTCTGTTGCAGTTGGCAGTTCTTTTCGGTTTTCCCTCCTGCTGTTTTGGGGGCA  
 TGAAAGGGCTTCGCCGCCGGGAGTAAAAGAAGGAATTGACCGGGCAGCGCGAGGGAGGAGCGCGCACGCGACCGC  
 GAGGGCGGGCGTGACCCCTCGGCTGGAAGTTTGTGCCGGGCCCCGAGCGCGCGCCGGCTGGGAGCTTCGGGTAGA  
 GACCTAGGCCGCTGGACCGCGATGAGCGCGCCGAGCCTCCGTGCGCGCGCCGCGGGGTGGGGCTGCTGCTGTGC  
 GCGGTGCTGGGGCGCGCTGGCCGGTCCGACAGCGCGGTGCGGGGAACTCGGGCAGCCCTCTGGGGTAGCCGCC  
 GAGCGCCCATGCCCACTACCTGCCGCTGCCTCGGGGACCTGCTGGACTGCAGTCGTAAGCGGCTAGCGCGTCTT  
 CCCGAGCCACTCCCGTCTGGGTGCTCGGCTGGACTTAAGTCACAACAGATTATCTTTCATCAAGGCAAGTTCC  
 ATGAGCCACCTTCAAAGCCTTCGAGAAGTGAAACTGAACAACAATGAATTGGAGACCATTCAAATCTGGGACCA  
 GTCTCGGCAAATATTACACTTCTCTCTTGGCTGGAAACAGGATTGTTGAAATACTCCCTGAACATCTGAAAGAG  
 TTTCACTCCCTTGAACCTTTGGACCTTAGCAGCAACAATATTTCAAGAGCTCCAACTGCATTTCCAGCCCTACAG  
 CTCAAATATCTGTATCTCAACAGCAACCGAGTCACATCAATGGAACCTGGGTATTTTGACAATTTGGCCAACACA  
 CTCCTTGTGTTAAAGCTGAACAGGAACCGAATCTCAGCTATCCCACCCAAGATGTTTAAACTGCCCAACTGCAA  
 CATCTCGAATTGAACCGAAACAAGATTAAAAATGTAGATGGACTGACATTTCCAAGGCCTTGGTGCTCTGAAGTCT  
 CTGAAAATGCAAAGAAATGGAGTAACGAACTTATGGATGGAGCTTTTTGGGGGCTGAGCAACATGGAAATTTTG  
 CAGCTGGACCATAACAACCTAACAGAGATTACCAAAGGCTGGCTTTACGGCTTGCTGATGCTGCAGGAACCTCAT  
 CTCAGCCAAAATGCCATCAACAGGATCAGCCCTGATGCCTGGGAGTTCTGCCAGAAGCTCAGTGAGCTGGACCTA  
 ACTTTCATCTGTTATCAAGGTTAGATTCAAGCTTCTTGGCCTAAGCTTACTAAATACACTGCACATTGGG  
 AACACAGAGTCAGCTACATTGCTGATTGTGCTTCCGGGGGCTTTCCAGTTTAAAGACTTTGGATCTGAAGAAC  
 AATGAAATTTCTGGACTATTGAAGACATGAATGGTGCTTCTCTGGGCTTGACAACTGAGGCGACTGATACTC  
 CAAGGAAATCGGATCCGTTCTATTACTAAAAAGCCTTCACTGGTTTGGATGCATTGGAGCATCTAGACCTGAGT  
 GACAACGCAATCATGTCTTACAAGGCAATGCATTTTCAAAATGAAGAAACTGCAACAATTGCATTTAAATACA  
 TCAAGCCTTTTGTGCGATTGCCAGCTAAAATGGCTCCACAGTGGGTGGCGGAAAACAACCTTTCAGAGCTTTGTA  
 AATGCCAGTTGTGCCCATCTCAGCTGCTAAAAGGAAGAAGCATTTTGTGCTGTAGCCAGATGGCTTTGTGTGT  
 GATGATTTTCCCAAACCCAGATCAGCGTTTCCAGCCAGAAACACAGTCGGCAATAAAAGGTTCCAATTTGAGTTT  
 ATCTGCTCAGCTGCCAGCAGCAGTGATTCCCAATGACTTTTGTCTGGAAAAAGACAATGAACACTGTCATGAT  
 GCTGAAATGGAAAAATATGCACACCTCCGGGCCCAAGGTGGCGAGGTGATGGAGTATACCACCATCCTTCGGCTG  
 CGCGAGGTGGAATTTGCCAGTGAGGGGAAATATCAGTGTGTCTCATCTCCAATCACTTTGGTTTCATCTACTCTGTC  
 AAAGCCAAGCTTACAGTAAATATGCTTCCCTCATTACCAAGACCCCCATGGATCTCACCATCCGAGCTGGGGCC  
 ATGGCACGCTTGGAGTGTGCTGTGTGGGGCACCCAGCCCCCAGATAGCCTGGCAGAAGGATGGGGGCACAGAC  
 TTCCAGCTGCACGGGAGAGACGCATGCATGTGATGCCCCGAGGATGACGTGTTCTTTATCGTGGATGTGAAGATA  
 GAGGACATTGGGGTATACAGCTGCACAGCTCAGAACAGTGCAGGAAGTATTTAGCAAATGCAACTCTGACTGTC  
 CTAGAAACACCATCATTTTTCGGGCCACTGTTGGACCGAACTGTAACCAAGGGAGAAACAGCCGCTCTACAGTGC  
 ATTTGCTGGAGGAAGCCCTCCCCCTAAACTGAACTGGACCAAGATGATAGCCCATTTGGTGGTAACCGAGAGGCAC  
 TTTTTTGCAGCAGGCAATCAGCTTCTGATTATTGTGGACTCAGATGTGATGCTGCGGAAATACACATGTGAG  
 ATGTCTAACACCCCTTGGCACTGAGAGAGGAACAGTCGCGCTCAGTGTGATCCCCACTCCAACCTGCGACTCCCC  
 CAGATGACAGCCCCATCGTTAGACGATGACGGATGGGCCACTGTGGGTGTCGTGATCATAGCCGTGGTTTGTGT  
 GTGGTGGGCACGTCACTCGTGTGGGTGGTCATCATATACCACACAAGGCGGAGGAATGAAGATTGCAGCATTACC  
 AACACAGATGAGACCAACTTGGCAGCAGATATTCTTAGTTATTTGTCTCATCTCAGGGAACGTTAGCTGACAGGCAG  
 GATGGGTACGTGTCTTCAGAAAGTGGAAGCCACCACAGTTTGTACATCTTCAGGTGCTGGATTTTTCTTACCA  
 CAACATGACAGTAGTGGGACCTGCCATATTGACAATAGCAGTGAAGCTGATGTGGAAGCTGCCACAGATCTGTTT  
 CTTTGTCCGTTTTTGGGATCCACAGGCCCTATGTATTTGAAGGAAATGTGTATGGCTCAGATCCTTTTGAACA  
 TATCATACAGGTTGCAGTCTTGACCCAAGAACAGTTTTAATGGACCACTATGAGCCAGTTACATAAAGAAAAAG  
 GAGTGCTACCCATGTTCTCATCCTTCAGAAGAATCCTGCGAACGGAGCTTCAGTAATATATCGTGGCCTTCACAT  
 GTGAGGAAGCTACTTAACACTAGTTACTCTCACAATGAAGGACCTGGAATGAAAAATCTGTGTCTAAACAAGTCC  
 TCTTTAGATTTTAGTGAAATCCAGAGCCAGCGTCGGTTGCCTCGAGTAATTCTTTCATGGGTACCTTTGGAAAA  
 GCTCTCAGGAGACCTCACCTAGATGCCTATTCAAGCTTTGGACGCCATCAGATTGTGACCCAAGAGCCTTTTAT  
 TTGAAAGCTCATTTCTCCCGAGCTTGGACTCTGGGTGAGGGAAGATGGGAAAGAAAGGACAGATTTTCAGGAA  
 GAAATCACATTTGTACCTTTAAACAGACTTTAGAAAACACAGGACTCCAAATTTTCAGTCTTATGACTTGGAC  
 ACATAGACTGAATGAGACCAAGGAAAAGCTTAACATACTACCTCAAGTGAACCTTTATTTAAAGAGAGAGAAT  
 CTTATGTTTTTTAAATGGAGTTATGAATTTTAAAGGATAAAAAATGCTTTATTTATACAGATGAACCAAAATTAC  
 AAAAGTTATGAAATTTTTATACTGGGAATGATGCTCATATAAGAATACCTTTTTAACTATTTTTTAACCTTG  
 TTTTATGCAAAAAAGTATCTTACGTAAATTAATGATATAAATCATGATTATTTTATGATTTTTTATAATGCCAGA  
 TTTCTTTTTATGGAAAATGAGTTACTAAAGCATTTTAAATAATACCTGCCTGTACCATTTTTTAAATAGAAGTT  
 ACTTCATTATATTTTGCACATTATATTTAATAAAATGTGTCAATTTGAAAAAAAAAAAAAAAAAAAAAAAAAAAA

## FIGURE 106

MSAPSLRARAAGLGLLLCAVLGRAGRSDSGRGELGQPSGVAAERPCPTTCRCLGDLLDCSR  
KRLARLPEPLPSWVARLDLSHNRLSFIKASSMSHLQSLREVKLNNNELETIPNLGPVSANIT  
LLSLAGNRIVEILPEHLKEFQSLETLDLSSNNISELQTAFPALQLKYLYLNSNRVTSMEPGY  
FDNLANTLLVLKLNRRNRI SAIPPKMFKL PQLOHLELNRNKIKNVDGLTFQGLGALKSLKMQR  
NGVTKLMDGAFWGLSNMEILQLDHNNLTEITKGWLYGLLMLQELHLSQNAINRISPDAWEFC  
QKLSELDLTFNHL SRLDDSSFLGLSLLNTLHIGNNRVSYIADCAFRGLSSLKTLDLKNNEIS  
WTIEDMNGAFSGLDKLRRILIQGNRIRSITKKAFTGLDALEHLDLSDNAIMSLQNAFSQMK  
KLQQLHLNTSSLLCDCQLKWL PQWVAENNFQSFVNASCAHPQLLKGRSIFAVSPDGFVCDDF  
PKPQITVQPETQSAIKGSNLSFICSAASSSDSPMTFAWKKNELLHDAEMENY AHLRAQGG  
VMEYTTILRLREVEFASEGKYQCVISNHFSSYSVKAKLTVNMLPSFTKTPMDLTIRAGAMA  
RLECAAVGHPAPQIAWQKDG GTDFPAARERRMHVMPEDDVFFIVDVKIEDIGVYSCTAQNSA  
GSISANATLTVLETSPFLRPLLDRTVTKGETAVLQCIAGGSPPPKNWTKDDSPLVVTERHF  
FAAGNQLLIIVDSDVSDAGKYTCEMSNTLGTERGNVRLSVIPTPTCDSPQMTAPSLDDDGWA  
TVGVVIIAVVCCVVGTSLVWVVI IYHTRRRNEDCSITNTDETNPADIPSYLSSQGT LADRQ  
DGYVSSESGSHHQFVTSSGAGFFLPQHDSSGTCHIDNSSEADVEAATDLFLCPFLGSTGPMY  
LKG NVYGS DPFETYHTGCSPDPRTVLMDHYEPSYIKKKECYPCSHPSEESCERSFSNISWPS  
HVRKLLNTSYSHNEGPGMKNLCLNKSSLD FSANPEPASVASSNSFMGTFGKALRRPHLDAYS  
SFGQPSDCQPRAFYLKAHSSPDLD SGSEEDGKERTDFQEENHICTFKQTLENYRTPNFQSYDLDT

**Signal sequence:**

amino acids 1-27

**Transmembrane domain:**

amino acids 808-828

**N-glycosylation site.**

amino acids 122-126, 156-160, 274-278, 442-446, 469-473, 515-519,  
688-692, 729-733, 905-909, 987-991, 999-1003, 1016-1020

**Glycosaminoglycan attachment site.**

amino acids 886-890

**Casein kinase II phosphorylation site.**

amino acids 99-103, 180-184, 263-267, 314-318, 324-328, 374-378,  
383-387, 407-411, 524-528, 608-612, 692-696, 709-713, 731-735,  
799-803, 843-847, 863-867, 907-911, 1003-1007, 1018-1022,  
1073-1077, 1079-1083, 1081-1085

**Tyrosine kinase phosphorylation site.**

amino acids 667-675

**N-myristoylation site.**

amino acids 14-20, 36-42, 239-245, 257-263, 380-386, 427-433,  
513-519, 588-594, 672-678, 683-687, 774-780, 933-939

**Leucine zipper pattern.**

amino acids 58-80, 65-87

# FIGURE 107

CAAAACCTTGCGTCGCGGAGAGCGCCAGCTTGACTTGAATGGAAGGAGCCCGAGCCCGCGGAGCGCAGCTGAGAC  
 TGGGGGAGCGCGTTTCGGCCTGTGGGGCGCCGCTCGGCGCCGGGGCGCAGCAGGGAAGGGGAAGCTGTGGTCTGCC  
 CTGCTCCACGAGGCGCCACTGGTGTGAACGGGAGAGCCCTGGGTGGTCCCGTCCCCTATCCCTCCTTTATATA  
 GAAACCTTCCACACTGGGAAGGCAGCGGCGAGGCAGGAGGGCTCATGGTGAAGGAGGCCGGCTGATCTGCAG  
 GCGCACAGCATTCAGATTTACAGATTTTACAGATACCAATGGAAGGCGAGGAGGCAGAACAGCCTGCCTGGT  
 TCCATCAGCCCTGGCGCCAGGCGCATCTGACTCGGCACCCCTGCAGGCACCATGGCCCGAGCCGGGTGCTGC  
 TGCTCCTGCTGCTGCTGCCGCCACAGCTGCACCTGGGACCTGTGCTTGCCGTGAGGGCCCCAGGATTTGGCCGAA  
 GTGGCGGCCACAGCCTGAGCCCCGAAGAGAACGAATTTGCGGAGGAGGAGCCGGTGTGGTACTGAGCCCTGAGG  
 AGCCCGGGCCTGGCCCGAGCCGCGGTGAGCTGCCCCGAGACTGTGCCTGTTCCAGGAGGGCGTCTGGACTGTG  
 GCGGTATTGACCTGCGTGAGTTCCCGGGGACCTGCCTGAGCACACCAACCACCTATCTCTGCAGAACCAACCAGC  
 TGGAAAAGATCTACCCTGAGGAGCTCTCCCGGCTGCACCGGTGGAGACACTGAACCTGCAAAACAACCGCCTGA  
 CTTCCCGAGGGCTCCCAGAGAAGGCGTTTGAGCATCTGACCAACCTCAATTACCTGTACTTGGCCAATAACAAGC  
 TGACCTTGGCACCCCGCTTCCCTGCCAAACGCCCTGATCAGTGTGGACTTTGCTGCCAACTATCTCACCAGATCT  
 ATGGGCTCACCTTTGGCCAGAAGCCAACTTGAGGTCTGTGTACCTGCACAACAACAAGCTGGCAGACGCCGGGC  
 TGCCGGACAACATGTTCAACGGCTCCAGCAACGTGAGGTCTCATCTGTCCAGCAACTTCTGCGCCACGTGC  
 CCAAGCACCTGCGCCTGCCCTGTACAAGCTGCACCTCAAGAACAACAAGCTGGAGAAGATCCCCCGGGGGCCT  
 TCAGCGAGCTGAGCAGCCTGCGCGAGCTATACCTGCAGAACAACTACCTGACTGACGAGGGCCTGGACAACGAGA  
 CCTTCTGGAAGCTCTCCAGCCTGGAGTACCTGGATCTGTCCAGCAACAACCTGTCTCGGGTCCCAGCTGGGCTGC  
 CGCGCAGCCTGGTGTGCTGCACTTGGAGAAGAAGCGCATCCGGAGCGTGGACGCGAATGTGCTGACCCCCATCC  
 GCAGCCTGGAGTACCTGCTGCTGCACAGCAACCAGCTGCGGGAGCAGGGCATCCACCCACTGGCCTTCCAGGGCC  
 TCAAGCGGTTGCACACGGTGCACCTGTACAACAACGCGCTGGAGCGCTGCCAGTGGCCTGCCTGCGCCGCTGC  
 GCACCTCATGATCCTGCACAACCAGATCAAGGCATTGGCCGCGAAGACTTTGCCACCACCTACTTCTGGAGG  
 AGCTCAACCTCAGCTACAACCGCATCACCAGCCACAGGTGCACCGCGACGCTTCCGCAAGCTGCGCCTGCTGC  
 GCTCGCTGGACCTGTGCGGCAACCGGCTGCACACGCTGCCACCTGGGCTGCCCTGAAATGTCCATGTGCTGAAGG  
 TCAAGCGCAATGAGCTGGCTGCCTTGGCACGAGGGGCGCTGGCGGGCATGGCTCAGCTGCGTGAGCTGTACCTCA  
 CCAGCAACCGACTGCGCAGCCGAGCCCTGGGCCCCCGTGCCTGGGTGGACCTCGCCCATCTGCAGCTGCTGGACA  
 TCGCCGGAATCAGCTCACAGAGATCCCCGAGGGGCTCCCCGAGTCACTTGAGTACCTGTACCTGCAGAACAAACA  
 AGATTAGTGCGGTGCCCCGCAATGCCTTCGACTCCACGCCCAACCTCAAGGGGATCTTTCTCAGGTTTAACAAGC  
 TGGCTGTGGGCTCCGTGGTGGACAGTGCCTTCCGGAGGCTGAAGCACCTGCAGGTCTTGGACATTGAAGGCAACT  
 TAGAGTTTGGTGACATTTCCAAGGACCTGGCCGCTTGGGGAAGGAAAAGGAGGAGGAGGAAGAGGAGGAGGAGG  
 AGGAAGAGGAAAACAAGATAGTGACAAGGTGATGCAGATGTGACCTAGGATGATGGACCGCCGACTCTTTCTGC  
 AGCACACGCTGTGTGCTGTGAGCCCCCACTCTGCCGTGCTCACACAGACACACCCAGCTGCACACATGAGGCA  
 TCCACATGACACGGGCTGACACAGTCTCATATCCCCACCCCTTCCCACGGCGTGTCCCACGGCCAGACACATGC  
 ACACACATCACACCTCAAACACCCAGCTCAGCCACACACAACCTACCTCCAAACCACACAGTCTCTGTACAC  
 CCCCCTACCGCTGCCACGCCCTCTGAATCATGCAGGGAAGGGTCTGCCCTGCCCTGGCACACACAGGCACCCA  
 TTCCCTCCCCCTGCTGACATGTGTATGCGTATGCATACACACACACACACATGCACAAGTCAATGTGCGAA  
 CAGCCCTCCAAAGCCTATGCCACAGACAGCTCTTGCCCCAGCCAGAATCAGCCATAGCAGCTCGCCGTCTGCCCT  
 GTCCATCTGTCCGTCCGTTCCTGGAGAAGACACAAGGGTATCCATGCTCTGTGGCCAGGTGCCTGCCACCCCTCT  
 GGAACCTCAAAAAGCTGGCTTTTATTCTTTCCCATCTATGGGGACAGGAGCCTTCAGGACTGCTGGCCTGGCC  
 TGGCCACCCCTGCTCCTCCAGGTGCTGGGCAGTCACTCTGCTAAGAGTCCCTCCCTGCCACGCCCTGGCAGGACA  
 CAGGCACTTTTCCAATGGGCAAGCCAGTGGAGGCAGGATGGGAGAGCCCCCTGGGTGCTGCTGGGGCCTTGGGG  
 CAGGAGTGAAGCAGAGGTGATGGGGCTGGGCTGAGCCAGGGAAGGACCCAGCTGCACCTAGGAGACACCTTT  
 GTTCTTCAGGCCTGTGGGGGAAGTTCCGGGTGCCTTTATTTTATTTCTTTCTAAGGAAAAAATGATAAAAT  
 CTCAAAGCTGATTTTCTTGTATAGAAAACTAATATAAAAGCATTATCCCTATCCCTGCAAAAAA

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
841  
842  
843  
844  
845  
846  
847  
848  
849  
850  
851  
852  
853  
854  
855  
856  
857  
858  
859  
860  
861  
862  
863  
864  
865  
866  
867  
868  
869  
870  
871  
872  
873  
874  
875  
876  
877  
878  
879  
880  
881  
882  
883  
884  
885  
886  
887  
888  
889  
890  
891  
892  
893  
894  
895  
896  
897  
898  
899  
900  
901  
902  
903  
904  
905  
906  
907  
908  
909  
910  
911  
912  
913  
914  
915  
916  
917  
918  
919  
920  
921  
922  
923  
924  
925  
926  
927  
928  
929  
930  
931  
932  
933  
934  
935  
936  
937  
938  
939  
940  
941  
942  
943  
944  
945  
946  
947  
948  
949  
950  
951  
952  
953  
954  
955  
956  
957  
958  
959  
960  
961  
962  
963  
964  
965  
966  
967  
968  
969  
970  
971  
972  
973  
974  
975  
976  
977  
978  
979  
980  
981  
982  
983  
984  
985  
986  
987  
988  
989  
990  
991  
992  
993  
994  
995  
996  
997  
998  
999  
1000

## **FIGURE 108**

MEGEEAEQPAWFHQWRPGASDSAPPAGTMAQSRVLLLLLLLLPPQLHLGPVLAVRAPGFGRS  
GGHSLSPREENEFAEEEPVLVLSPPEPGPGPAAVSCPRDCACSQEGVVDCCGIDLREFPGDLP  
EHTNHLSQLNNQLEKIYPEELSRHLRLETNLQNNRLTSRGLPEKAFEHLTNLNYLYLANNK  
LTLAPRFLPNALISVDFAANYLTKIYGLTFGQKPNLRSVYLHNNKLADAGLPDNMFNGSSNV  
EVLILSSNFLRHVPKHLPPALYKLHLKNNKLEKIPPGAFSELSSLRELYLQNNYLTDEGLDN  
ETFWKLSSLEYLDLSSNNLSRVPAGLPRSLVLLHLEKNAIRSVDANVLTPIRSLEYLLHLSN  
QLREQGIHPLAFQGLKRLHTVHLYNNALERVPSGLPRRVRTLMLHNQITGIGREDFATTYF  
LEELNLSYNRITSPQVHRDAFRKLRLRLSLDLSGNRLHTLPPGLPRNVHVLKVKRNELAALA  
RGALAGMAQLRELYLTSNRLRSRALGPRAWVDLAHLQLLDIAGNQLTEIPEGLPESLEYLYL  
QNNKISAVPANAFDSTPNLKGIFLRFNKLAVGSVVDSAFRRRLKHLQVLDIEGNLEFGDISKD  
RGRLGKEKEEEEEEEEEEEEEETR

### **Signal sequence:**

amino acids 1-48

### **N-glycosylation site.**

amino acids 243-247, 310-314, 328-332, 439-443

### **Casein kinase II phosphorylation site.**

amino acids 68-72, 84-88, 246-250, 292-296, 317-321, 591-595

### **N-myristoylation site.**

amino acids 19-25, 107-113, 213-219, 217-223, 236-242, 335-341,  
477-483, 498-502, 539-545, 548-554

### **Leucine zipper pattern.**

amino acids 116-138, 251-273, 258-280, 322-344, 464-486, 471-493,  
535-557

biochem. biophys. Res. Commun.



## FIGURE 109

GGGAGGGGGCTCCGGGCGCCGCGCAGCAGACCTGCTCCGGCCGCGCGCCTCGCCGCTGTCTCCGGGAGCGGCAG  
CAGTAGCCCCGGGCGGCGAGGGCTGGGGGTTCCTCGAGACTCTCAGAGGGGCGCCTCCCATCGGCGCCCCACCACC  
CAACCTGTTCTCGCGCGCCACTGCGCTGCGCCCCAGGACCCGCTGCCAACATGGATTTTCTCCTGGCGCTGGT  
GCTGGTATCCTCGCTCTACCTGCAGGCGGCGCCGAGTTCGACGGGAGGTGGCCAGGCAAATAGTGTATCGAT  
TGGCCTATGTCGTTATGGTGGGAGGATTGACTGCTGCTGGGCTGGGCTCGCCAGTCTTGGGGACAGTGTAGCC  
TGTGTGCCAACCCAGATGCAACATGGTGAATGTATCGGGCCAAACAAGTGCAAGTGTATCCTGGTTATGCTGG  
AAAAACCTGTAATCAAGATCTAAATGAGTGTGGCCTGAAGCCCCGGCCCTGTAAGCACAGGTGCATGAACACTTA  
CGGAGCTACAAGTGTACTGTCTCAACGGATATATGCTCATGCCGATGGTTCCTGCTCAAGTGCCCTGACCTG  
CTCCATGGCAAACCTGTAGTATGGCTGTGATGTTGTTAAAGGACAAATACGGTGCCAGTGCCCATCCCCCTGGCCT  
GCACCTGGCTCCTGATGGGAGGACCTGTGTAGATGTTGATGAATGTGCTACAGGAAGAGCCTCCTGCCCTAGATT  
TAGGCAATGTGTCAACACTTTTGGGAGCTACATCTGCAAGTGTATAAAGGCTTCGATCTCATGTATATTGGAGG  
CAAATATCAATGTATGACATAGACGAATGCTCACTGGTCAGTATCAGTGCAGCAGCTTTGCTCGATGTTATAA  
CGTACGTGGGTCTTACAAGTGCAATGTAAAGAAGGATACCAGGGTGATGGACTGACTTGTGTGTATATCCCAAA  
AGTTATGATTGAACCTTCAGGTCCAATTCATGTACCAAGGGAAATGGTACCATTTTAAAGGGTGACACAGGAAA  
TAATAATTGGATTCCCTGATGTTGGAAGTACTTGGTGGCCTCCGAAGACACCATATATTCCTCCTATCATTACCAA  
CAGGCCTACTTCTAAGCCAACAACAAGACCTACACCAAGCCAACACCAATTCCTACTCCACCACCACCACCACC  
CCTGCCAACAGAGCTCAGAACACCTCTACCACCTACAACCCAGAAAGGCCAACACCACCGGACTGACAACTATAGC  
ACCAGCTGCCAGTACACCTCCAGGAGGGATTACAGTTGACAACAGGGTACAGACAGACCCTCAGAAACCCAGAGG  
AGATGTGTTTCACTGTTCTGGTACACAGTTGTAATTTTGACCATGGACTTTGTGGATGGATCAGGGAGAAAGACAA  
TGACTTGCACTGGGAACCAATCAGGGACCCAGCAGGTGGACAATATCTGACAGTGTCCGCAGCCAAAGCCCCAGG  
GGGAAAAGCTGCACGCTTGGTGCTACCTCTCGGCCGCGCTCATGCATTAGGGGACCTGTGCCTGTCAATCAGGCA  
CAAGGTGACGGGGCTGCACTCTGGCACACTCCAGGTGTTTGTGAGAAAACACGGTGCCACGGAGCAGCCCTGTG  
GGGAAGAAATGGTGGCCATGGCTGGAGGCAACACAGATCACCTTGCAGGGGGCTGACATCAAGAGCGAATCACA  
AAGATGATTAAAGGGTTGGAAAAAAGATCTATGATGGAAAATTAAGGAACTGGGATTATTGAGCCTGGAGAAG  
AGAAGACTGAGGGGCAAACCATTTGATGGTTTTCAAGTATATGAAGGGTTGGCACAGAGAGGGTGGCGACCAGCTG  
TTCTCCATATGCACTAAGAATAGAACAAGAGGAACTGGCTTAGACTAGAGTATAAGGGAGCATTTCCTGGCAGG  
GGCCATTGTTAGAATACTTCATAAAAAAAGAAGTGTGAAAATCTCAGTATCTCTCTCTCTTCTTCTAAAAAATTAGA  
TAAAAATTTGTCTATTTAAGATGGTTAAAGATGTTCTTACCCAAGGAAAAAGTAACAAATTATAGAATTTCCCAAA  
AGATGTTTTGATCCTACTAGTAGTATGCACTGAAAATCTTAGAACTAAATAATTTGGACAAGGCTTAATTTAGG  
CATTTCCCTCTTGACCTCCTAATGGAGAGGGATTGAAAGGGGAAGAGCCCACCAAATGCTGAGCTCACTGAAATA  
TCTCTCCCTTATGGCAATCCTAGCAGTATTAAAGAAAAAAGGAACTATTTATCCAAATGAGAGTATGATGGAC  
AGATATTTTAGTATCTCAGTAATGTCCTAGTGTGGCGGTGGTTTTCAATGTTTCTTCATGGTAAAGGTATAAGCC  
TTTCATTTGTTCAATGGATGATGTTTCAGATTTTTTTTTTTTTTAAAGAGATCCTTCAAGGAACACAGTTTCAGAGAG  
ATTTTCATCGGGTGCACTCTCTCTGCTTCTGTGTGACAAGTTATCTTGGCTGCTGAGAAAGAGTGCCCTGCCCC  
ACACCGGCAGACCTTTCTTCACCTCATCAGTATGATTTCAGTTTCTCTTATCAATTGGACTCTCCCAGGTTCCAC  
AGAACAGTAATATTTTTTGAACCAATAGGTACAATAGAAGGTCTTCTGTCAATTAACCTGGTAAAGGCAGGGCTGG  
AGGGGGAAAATAAATCATTAAGCCTTTGAGTAACGGCAGAATATATGGCTGTAGATCCATTTTAAATGGTTTCATT  
TCCTTTATGGTCATATAACTGCACAGCTGAAGATGAAAGGGGAAAATAAATGAAAATTTTACTTTTCGATGCCAA  
TGATACATTGCACTAAACTGATGGAAGAAGTTATCCAAAGTACTGTATAACATCTTGTATTATTATTTAATGTTTT  
CTAAAATAAAAAATGTTAGTGGTTTTCCAAATGGCCTAATAAAAAACAATTATTTGTAAATAAAAAACACTGTTAGTAAT

## **FIGURE 110**

MDFLLALVLVSSSLYLQAAAEFDGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPVCQP  
RCKHGEICIGPNKCKCHPGYAGKTCNQDLNECGLKPRPCKHRCMNTYGSYKCYCLNGYMLMPD  
GSCSSALTCSMANCQYGC DVVKGQIRCQCPSPLHLAPDGRTCDVDVDECATGRASCPRFRQC  
VNTFGSYICKCHKGFDLMYIGGKYQCHDIDEC SLGQYQCSSFARCYNVRGSYKCKCKEGYQG  
DGLTCVYIPKVMIEPSGPIHV PKGNGTILKGDTGNNNWIPDVGSTWWPPKTPYIPPIITNRP  
TSKPTTRPTPKPTPIPTPPPPPLPTELRTPLPPTTPTTGLTTIAPAASTPPGGITVDN  
RVQTD PQKPRGDVFSVLVHSCNFDHGLCGWIREKDNDLHWEPIRDPAGGQYLTVSAAKAPGG  
KAARLVLP LGRMLHSGDLCLSFRHKVTGLHSGTLQVFVRKHGAHGAALWGRNGGHGWRQTQI  
TLRGADIKSESQR

### **Signal sequence:**

amino acids 1-17

### **N-glycosylation site.**

amino acids 273-277

### **Casein kinase II phosphorylation site.**

amino acids 166-170, 345-349

### **Tyrosine kinase phosphorylation site.**

amino acids 199-206

### **N-myristoylation site.**

amino acids 109-115, 125-131, 147-153, 191-197, 221-227, 236-242,  
421-427, 433-439, 462-468, 476-482

### **Aspartic acid and asparagine hydroxylation site.**

amino acids 104-116, 186-198, 231-243

### **Cell attachment sequence.**

amino acids 382-385

### **EGF-like domain cysteine pattern signature.**

amino acids 75-87

## FIGURE 111

CTTCTTTGAAAAGGATTATCACCTGATCAGGTTCTCTCTGCATTGCCCCCTTTAGATTGTGA  
AATGTGGCTCAAGGTCTTCACAACCTTTCTTTCTTTGCAACAGGTGCTTGCTCGGGGCTGA  
AGGTGACAGTGCCATCACACACTGTCCATGGCGTCAGAGGTCAGGCCCTCTACCTACCCGTC  
CACTATGGCTTCCACACTCCAGCATCAGACATCCAGATCATATGGCTATTTGAGAGACCCCA  
CACAATGCCCAAATACTTACTGGGCTCTGTGAATAAGTCTGTGGTTCCTGACTTGGAATACC  
AACACAAGTTTACCATGATGCCACCCAATGCATCTCTGCTTATCAACCCACTGCAGTTCCCT  
GATGAAGGCAATTACATCGTGAAGGTCAACATTGAGGGAAATGGAAGTCTATCTGCCAGTCA  
GAAGATACAAGTCACGGTTGATGATCCTGTCAAAAGCCAGTGGTGCAGATTCATCCTCCCT  
CTGGGGCTGTGGAGTATGTGGGGAACATGACCTGACATGCCATGTGGAAGGGGGCACTCGG  
CTAGCTTACCAATGGCTAAAAAATGGGAGACCTGTCCACACCAGCTCCACCTACTCCTTTTC  
TCCCCAAAACAATAACCCTTCATATTGCTCCAGTAACCAAGGAAGACATTGGGAATTACAGCT  
GCCTGGTGAGGAACCCTGTCAGTGAAATGGAAAGTGATATCATTATGCCCATCATATATTAT  
GGACCTTATGGACTTCAAGTGAATTCTGATAAAGGGCTAAAAGTAGGGGAAGTGTTTACTGT  
TGACCTTGGAGAGGCCATCCTATTTGATTGTTCTGCTGATTCTCATCCCCCAACACCTACT  
CCTGGATTAGGAGGACTGACAATACTACATATATCATTAAGCATGGGCCTCGCTTAGAAGTT  
GCATCTGAGAAAGTAGCCCAAGAACAATGGACTATGTGTGCTGTGCTTACAACAACATAAC  
CGGCAGGCAAGATGAAACTCATTTACAGTTATCATCACTTCCGTAGGACTGGAGAAGCTTG  
CACAGAAAGGAAAATCATTGTACCTTTAGCAAGTATAACTGGAATATCACTATTTTTTGATT  
ATATCCATGTGTCTTCTCTTCTATGGAAAAAATATCAACCCTACAAAGTTATAAAACAGAA  
ACTAGAAGGCAGGCCAGAAACAGAATACAGGAAAGCTCAAACATTTTCAGGCCATGAAGATG  
CTCTGGATGACTTCGGAATATATGAATTTGTTGCTTTTCCAGATGTTTCTGGTGTTTCCAGG  
ATTCCAAGCAGGTCTGTTCCAGCCTCTGATTGTGTATCGGGGCAAGATTTGCACAGTACAGT  
GTATGAAGTTATTCAGCACATCCCTGCCAGCAGCAAGACCATCCAGAGTGAACCTTTCATGG  
GCTAAACAGTACATTCGAGTGAAATTCTGAAGAAACATTTTAAGGAAAAACAGTGGAAAAGT  
ATATTAATCTGGAATCAGTGAAGAAACCAGGACCAACACCTCTTACTCATTATTCCTTTACA  
TGCAGAATAGAGGCATTTATGCAAATTGAACTGCAGGTTTTTCAGCATATACACAATGTCTT  
GTGCAACAGAAAAACATGTTGGGGAAATATTCCTCAGTGGAGAGTCGTTCTCATGCTGACGG  
GGAGAACGAAAGTGACAGGGGTTTCTCATAAGTTTTGTATGAAATATCTCTACAAACCTCA  
ATTAGTTCTACTCTACACTTTCCTATCATCAACACTGAGACTATCCTGTCTCACCTACAAA  
TGTGGAAACTTTACATTGTTTCGATTTTTTCAGCAGACTTTGTTTTATTAAATTTTTATTAGTG  
TTAAGAATGCTAAATTTATGTTTCAATTTTTATTTCCAAATTTCTATCTTGTTATTTGTACAA  
CAAAGTAATAAGGATGGTTGTCAAAAAACAAAACCTATGCCTTCTCTTTTTTTTTCAATCACC  
AGTAGTATTTTTTGAGAAGACTTGTGAACACTTAAGGAAATGACTATTAAAGTCTTATTTTTA  
TTTTTTTTCAAGGAAAGATGGATTCAAATAAATTATTCTGTTTTTGCTTTTAAAAAAAAAAAAA

## **FIGURE 112**

MWLKVFTTFLSFATGACSGLKVTVPSTVHGVRGQALYLPVHYGFHTPASDIQIIWLFERPH  
TMPKYLLGSVNKSVPDLEYQHKFTMMPPNASLLINPLQFPDEGNYIVKVNIQNGTLSASQ  
KIQVTVDDPVTKPVVQIHPPSGAVEYVGNMTLTCHVEGGTRLAYQWLKNGRPVHTSSTYSFS  
PQNNTLHIAPVTKEDIGNYSCLVRNPVSEMESDIIMPIIYYGPYGLQVNSDKGLKVGEVFTV  
DLGEAILFDCSADSHPPNTYSWIRRTDNTTYIIKHGPRLEVASEKVAQKTMDYVCCAYNNIT  
GRQDETHFTVITITSVGLEKLAQKGKSLSPASITGISLFLIISMCLLFLWKKYQPYKVIKQK  
LEGRPETEYRKAQTFSGHEDALDDFGIYEFVAFPDVSGVSRIPSRVSPASDCVSGQDLHSTV  
YEVIQHIPAQQDDHPE

### **Signal sequence:**

amino acids 1-18

### **Transmembrane domain:**

amino acids 341-359

### **N-glycosylation site.**

amino acids 73-77, 92-96, 117-121, 153-157, 189-193, 204-208,  
276-280, 308-312

### **Casein kinase II phosphorylation site.**

amino acids 129-133, 198-202, 214-218, 388-392, 426-430, 433-437

### **Tyrosine kinase phosphorylation site.**

amino acids 272-280

### **N-myristoylation site.**

amino acids 15-21, 19-25, 118-124, 163-167, 203-209, 231-237,  
239-245

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 7-18

1. The first step is to identify the key components of the system. This involves understanding the hardware, software, and data involved.

GCAAGCGCGCGAAATCGCGCCCTCCGGGAGTCTTGACAGTTCCCTTGGCAGTCTCTGGTGCTGT  
GCTTTGGGGTGCTCCCTGGACGCACGGGCGGCGGAGCAACGTTTCGCGTCATCACGGACGAGA  
ACTGGAGAGAACTGCTGGAAGGAGACTGGATGATAGAATTTTATGCCCCGTGGTGCCCTGCT  
TGTCAAAATCTTCAACCGGAATGGGAAAGTTTTGCTGAATGGGGAGAAGATCTTGAGGTTAA  
TATTGCGAAAGTAGATGTCACAGAGCAGCCAGGACTGAGTGGACGGTTTATCATAACTGCTC  
TTCCTACTATTTATCATTGTAAAGATGGTGAATTTAGGCGCTATCAGGGTCCAAGGACTAAG  
AAGGACTTCATAAACTTTATAAGTGATAAAGAGTGGAAGAGTATTGAGCCCGTTTCATCATC  
GTTTGGTCCAGGTTCTGTTCTGATGAGTAGTATGTCAGCACTCTTTCAGCTATCTATGTGGA  
TCAGGACGTGCCATAACTACTTTATTGAAGACCTTGGATTGCCAGTGTGGGGATCATATACT  
GTTTTTGTCTTTAGCAACTCTGTTTTCCGGACTGTTATTAGGACTCTGTATGATATTTGTGGC  
AGATTGCCTTTGTCTTCAAAAAGGCGCAGACCACAGCCATAACCATAACCCTTCAAAAAAAT  
TATTATCAGAATCTGCACAACCTTTGAAAAAAGTGAGAGGAGGAACAAGAGGCGGATGAAGAA  
GATGTTTCAGAAGAAGAAGCTGAAAGTAAAGAAGGAACAAACAAAGACTTTCCACAGAATGC  
CATAAGACAACGCTCTCTGGGTCCATCATTGGCCACAGATAAATCCTAGTTAAATTTTATAG  
TTATCTTAATATTATGATTTTGATAAAACAGAAGATTGATCATTTTGTTTGGTTTGAAGTG  
AACTGTGACTTTTTTGAATATTGCAGGGTTCAGTCTAGATTGTCAATAATTGAAGAGTCTA  
CATTGAGAACATAAAAGCACTAGGTATACAAGTTTGAAATATGATTTAAGCACAGTATGATG  
GTTTAAATAGTTCTCTAATTTTTGAAAAATCGTGCCAAGCAATAAGATTTATGTATATTTGT  
TTAATAATAACCTATTTCAAGTCTGAGTTTTGAAAATTTACATTTCCCAAGTATTGCATTAT  
TGAGGTATTTAAGAAGATTATTTTAGAGAAAAATATTTCTCATTTGATATAATTTTTCTCTG  
TTTCACTGTGTGAAAAAAGAAGATATTTCCCATAAATGGGAAGTTTGCCCATTTGTCTCAAG  
AAATGTGTATTTTCAGTGACAATTTTCGTGGTCTTTTTAGAGGTATATTCCAAAATTTCTTGT  
ATTTTTAGGTTATGCAACTAATAAAAACTACCTTACATTAATTAATTACAGTTTTCTACACA  
TGGTAATACAGGATATGCTACTGATTTAGGAAGTTTTTAAAGTTCATGGTATTCTCTTGATT  
CAACAAAGTTTGATTTTCTCTTGATTTTTTCTTACTTACTATGGGTACATTTTTTTATTTTT  
CAAATTGGATGATAATTTCTTGGAACATTTTTTATGTTTTAGTAAACAGTATTTTTTTTGT  
GTTTCAAACCTGAAGTTTACTGAGAGATCCATCAAATTGAACAATCTGTTGTAATTTAAAT  
TTGGCCACTTTTTTCAGATTTTACATCATTCTTGCTGAACTTCAACTTGAAATTGTTTTTT  
TTTCTTTTTGGATGTGAAGGTGAACATTCTTGATTTTTGTCTGATGTGAAAAAGCCTTGGTA  
TTTTACATTTTGAAATTTCAAAGAAGCTTAATATAAAAGTTTGCATTCTACTCAGGAAAAAG  
CATCTTCTTGATATATGCTTAAATGTATTTTTGTCTCATATACAGAAAGTTCTTAATTGAT  
TTTACAGTCTGTAATGCTTGATGTTTTAAATAATAACATTTTTTATATTTTTTAAAGACAA  
ACTTCATGATTATCCTGTGTTCTTCTGACTGGTAATATTGTGTGGGATTTACAGGTA  
GTACAGTAGGATGGAACATTTTAGTGATTTTTTACTCCTTAAAGAGCTAGAATACATAGTTTT  
CACCTTAAAGAAGGGGGAAAAATCATAAATCAATGAATGAACACTGACCATACGTAGTAGAC  
AATTTCTGTAATGTCCCTTCTTTCTAGGCTCTGTTGCTGTGTGAATCCATTAGATTTACAG  
TATCGTAATATACAAGTTTTCTTTAAAGCCCTCTCCTTTAGAATTTAAATATTGTACCATT  
AAAGAGTTTGGATGTGTAACCTTGTGATGCCTTAGAAAAATATCCTAAGCACAAAATAAACCT  
TTCTAACCACTTCATTAAGCTGAAAAAAAAAAAAAAAAAAAA

## **FIGURE 114**

MAPSGSLAVPLAVLVLLLWGAPWTHGRRSNVRVITDENWRELLEGDWMIEFYAPWCPACQNL  
QPEWESFAEWGEDLEVNIKVVDVTEQPGLSGRFIITALPTIYHCKDGEFRRYQGPRTKKDFI  
NFISDKEWKSIEPVSSWFGPGSVLMSSMSALFQLSMWIRTCHNYFIEDLGLPVWGSYTVFAL  
ATLFSGLLLGLCMIFVADCLCPSKRRRPQPYPYPSKKLLSESAQPLKKVEEQEAEDEEDVSE  
EEAESKEGTNKDFPQNPAIRQ RSLGPSLATDKS

### **Signal sequence:**

amino acids 1-26

### **Transmembrane domain:**

amino acids 182-201

### **Casein kinase II phosphorylation site.**

amino acids 68-72, 119-123, 128-132, 247-251, 257-261

### **Tyrosine kinase phosphorylation site.**

amino acids 107-115

### **N-myristoylation site.**

amino acids 20-26, 192-198

### **Amidation site.**

amino acids 25-29

## FIGURE 115

GCGAGTGTCCAGCTGCGGAGACCCGTGATAATTCTGTTAACTAATTCAACAAACGGGACCCTT  
CTGTGTGCCAGAAACCGCAAGCAGTTGCTAACCAGTGGGACAGGCGGATTGGAAGAGCGGG  
AAGGTCCTGGCCCAGAGCAGTGTGACACTTCCCTCTGTGACCATGAAACTCTGGGTGTCTGC  
ATTGCTGATGGCCTGGTTTGGTGTCTGAGCTGTGTGCAGGCCGAATTCTTCACCTCTATTG  
GGCACATGACTGACCTGATTTATGCAGAGAAAGAGCTGGTGCAGTCTCTGAAAGAGTACATC  
CTTGTGGAGGAAGCCAAGCTTTCCAAGATTAAGAGCTGGGCCAACAAAATGGAAGCCTTGAC  
TAGCAAGTCAGCTGCTGATGCTGAGGGCTACCTGGCTCACCTGTGAATGCCTACAACTGG  
TGAAGCGGCTAAACACAGACTGGCCTGCGCTGGAGGACCTTGTCTGCAGGACTCAGCTGCA  
GGTTTTATCGCCAACCTCTCTGTGCAGCGGCAGTTCTTCCCCACTGATGAGGACGAGATAGG  
AGCTGCCAAAGCCCTGATGAGACTTCAGGACACATACAGGCTGGACCCAGGCACAATTTCCA  
GAGGGGAACCTTCAGGAACCAAGTACCAGGCAATGCTGAGTGTGGATGACTGCTTTGGGATG  
GGCCGCTCGGCCTACAATGAAGGGGACTATTATCATAACGGTGTGTGGATGGAGCAGGTGCT  
AAAGCAGCTTGATGCCGGGGAGGAGGCCACCACAACCAAGTCACAGGTGCTGGACTACCTCA  
GCTATGCTGTCTTCCAGTTGGGTGATCTGCACCGTGCCCTGGAGCTCACCCGCCGCCTGCTC  
TCCCTTGACCCAAGCCACGAACGAGCTGGAGGGGAATCTGCGGTACTTTGAGCAGTTATTGGA  
GGAAGAGAGAGAAAAACGTTAAACAAATCAGACAGAAGCTGAGCTAGCAACCCAGAGGCA  
TCTATGAGAGGCCTGTGGACTACCTGCCTGAGAGGGATGTTTACGAGAGCCTCTGTCTGGG  
GAGGGTGTCAAACCTGACACCCCGTAGACAGAAGAGGCTTTTCTGTAGGTACCACCATGGCAA  
CAGGGCCCCACAGCTGCTCATTGCCCCCTTCAAAGAGGAGGACGAGTGGGACAGCCCGCACA  
TCGTCAGGTACTACGATGTCATGTCTGATGAGGAAATCGAGAGGATCAAGGAGATCGCAAAA  
CCTAAACTTGCACGAGCCACCGTTCGTGATCCCAAGACAGGAGTCCTCACTGTGCCAGCTA  
CCGGGTTTCCAAAAGCTCCTGGCTAGAGGAAGATGATGACCCTGTTGTGGCCCGAGTAAATC  
GTCGGATGCAGCATATCACAGGGTTAACAGTAAAGACTGCAGAATTGTTACAGGTGCAAAT  
TATGGAGTGGGAGGACAGTATGAACCGCACTTCGACTTCTCTAGGCGACCTTTTGACAGCGG  
CCTCAAAACAGAGGGGAATAGGTTAGCGACGTTTCTTAACTACATGAGTGATGTAGAAGCTG  
GTGGTGCCACCGTCTTCCCTGATCTGGGGGCTGCAATTTGGCCTAAGAAGGGTACAGCTGTG  
TTCTGGTACAACCTCTTGCGGAGCGGGGAAGGTGACTACCGAACAAGACATGCTGCCTGCCC  
TGTGCTTGTGGGCTGCAAGTGGGTCTCCAATAAGTGGTTCCATGAACGAGGACAGGAGTTCT  
TGAGACCTTGTGGATCAACAGAAGTTGACTGAATCCTTTTCTGTCTTCCCCCTTCCTGGTC  
CTTCAGCCCATGTCAACGTGACAGACACCTTTGTATGTTCTTTGTATGTTCTTATCAGGCT  
GATTTTTGGAGAAATGAATGTTTGTCTGGAGCAGAGGGAGACCATACTAGGGCGACTCCTGT  
GTGACTGAAGTCCCAGCCCTTCATTTCAGCCTGTGCCATCCCTGGCCCCAAGGCTAGGATCA  
AAGTGGCTGCAGCAGAGTTAGCTGTCTAGCGCCTAGCAAGGTGCCCTTTGTACCTCAGGTGTT  
TTAGGTGTGAGATGTTTCAGTGAACCAAAGTTCTGATACCTTGTTTACATGTTTGTGTTTTAT  
GGCATTCTATCTATTGTGGCTTTACCAAAAAATAAAATGTCCCTACCAGAAAAA

[illegible]

**Signal sequence:**

N-glycosylation site.

**Glycosaminoglycan attachment site.**

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

**Casein kinase II phosphorylation site.**

**Tyrosine kinase phosphorylation site.**

amino acids 71-80, 489-496

N-myristoylation site.

amino acids 14-20, 131-137, 171-177, 446-452

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 8-19

Leucine zipper pattern.

amino acids 213-235



## FIGURE 117

GCAGTATTGAGTTTTACTTCCTCCTCTTTTGTAGTGAAGACAGACCATAATCCCAGTGTGAGTGAAATTGATTGT  
TTCATTTATTACCGTTTTGGCTGGGGGTAGTTCCGACACCTTCACAGTTGAAGAGCAGGCAGAAGGAGTTGTGA  
AGACAGGACAATCTTCTTGGGGATGCTGGTCCCTGGAAGCCAGCGGGCCTTGCTCTGTCTTTGGCCTCATTGACCC  
CAGGTTCTCTGGTTAAACTGAAAGCCTACTACTGGCCTGGTGCCCATCAATCCATTGATCCTTGAGGCTGTGCC  
CCTGGGGCACCCACCTGGCAGGGCCTACCACCATGCGACTGAGCTCCCTGTTGGCTCTGCTGCGGCCAGCGCTTC  
CCCTCATCTTAGGGCTGTCTCTGGGGTGACGCTGAGCCTCCTGCGGGTTTCTGGATCCAGGGGGAGGGAGAAG  
ATCCCTGTGTGAGGCTGTAGGGGAGCGAGGAGGGCCACAGAATCCAGATTGAGAGCTCGGCTAGACCAAAGTG  
ATGAAGACTTCAAACCCCGGATTGTCCCCTACTACAGGGACCCCAACAAGCCCTACAAGAAGGTGCTCAGGACTC  
GGTACATCCAGACAGAGCTGGGCTCCCGTGAGCGGTTGCTGGTGGCTGTCTGACCTCCCGAGCTACACTGTCCA  
CTTTGGCCGTGGCTGTGAACCGTACGGTGCGCCATCACTTCCCTCGGTTACTCTACTTCACTGGGCAGCGGGGG  
CCCGGGCTCCAGCAGGGATGAGGTTGGTGTCTCATGGGGATGAGCGGCCCGCCTGGCTCATGTCAGAGACCCTGC  
GCCACCTTCACACACACTTTGGGGCCGACTACGACTGGTTCTTCATCATGCAGGATGACACATATGTGCAGGCCC  
CCCGCTGGCAGCCCTTGCTGGCCACCTCAGCATCAACCAAGACCTGTACTTAGGCCGGGCAGAGGAGTTTATTG  
GCGCAGGCGAGCAGGCCCGGTACTGTCTATGGGGGCTTTGGCTACCTGTTGTACGGAGTCTCCTGCTTCGTCTGC  
GGCCACATCTGGATGGCTGCCGAGGAGACATTCTCAGTGCCCGTCTGACGAGTGCGCTTGGACGCTGCCTCATTG  
ACTCTCTGGGCGTGGCTGTGTCTCAGCAGCCAGGGGCAGCAGTATCGCTCATTGAACTGGCCAAAAATAGGG  
ACCCTGAGAAGGAAGGGAGCTCGGCTTTCCTGAGTGCCCTTCGCCGTGCACCCCTGTCTCCGAAGGTACCCTCATGT  
ACCGGCTCCACAAACGCTTCAGCGCTCTGGAGTTGGAGCGGGCTTACAGTGAAATAGAACAACTGCAGGCTCAGA  
TCCGGAACCTGACCGTGCTGACCCCCGAAGGGGAGGCAGGGCTGAGCTGGCCCGTTGGGCTCCCTGCTCCTTTCA  
CACCACACTCTCGCTTTGAGGTGCTGGGCTGGGACTACTTCACAGAGCAGCACACCTTCTCCTGTGCAGATGGGG  
CTCCCAAGTGCCCACTACAGGGGGCTAGCAGGGCGGACGTGGGTGATGCGTTGGAGACTGCCCTGGAGCAGCTCA  
ATCGGCGCTATCAGCCCCGCTGCGCTTCCAGAAGCAGCGACTGCTCAACGGCTATCGGCGCTTCGACCCAGCAC  
GGGGCATGGAGTACACCTGGACCTGCTGTTGGAATGTGTGACACAGCGTGGGCACCGGCGGGGCCCTGGCTCGCA  
GGGTGAGCTGCTGCGGCCACTGAGCCGGGTGGAATCCTACCTATGCCCTATGTCACTGAGGCCACCCGAGTGC  
AGCTGGTGCTGCCACTCCTGGTGGCTGAAGCTGCTGCAGCCCCGGCTTTCTCGAGGCGTTTGAGCCCAATGTCC  
TGGAGCCACGAGAACATGCATTGCTCACCCTGTTGCTGGTCTACGGGCCACGAGAAGGTGGCCGTGGAGCTCCAG  
ACCCATTTCTTGGGGTGAAGGCTGCAGCAGCGGAGTTAGAGCGACGGTACCCTGGGACGAGGCTGGCCTGGCTCG  
CTGTGCGAGCAGAGGCCCTTCCAGGTGCGACTCATGGACGTGGTCTCGAAGAAGCACCCCTGTGGACACTCTCT  
TCTTCCTTACCACCGTGTGGACAAGGCCTGGGCCCCGAAGTCTCAACCGCTGTGCGATGAATGCCATCTCTGGCT  
GGCAGGCCTTCTTCCAGTCCATTTCCAGGAGTTCAATCCTGCCCTGTCAACACAGAGATCACCCCCAGGGCCCC  
CGGGGGCTGGCCCTGACCCCCCTCCCCCTCCTGGTGCTGACCCCTCCCGGGGGGCTCCTATAGGGGGGAGATTTG  
ACCGGCAGGCTTCTGCGGAGGGCTGCTTCTACAACGCTGACTACCTGGCGGCCCGAGCCCGGCTGGCAGGTGAAC  
TGGCAGGCCAGGAAGAGGAGGAAGCCCTGGAGGGGCTGGAGGTGATGGATGTTTCTCCTCGGTTCTCAGGGCTCC  
ACCTCTTTCGGGCGGTAGAGCCAGGGCTGGTGCAAGATTCTCCCTGCGAGACTGCAGCCCACGGCTCAGTGAAG  
AACTCTACCACCGCTGCCGCCTCAGCAACCTGGAGGGGCTAGGGGGCCGTGCCAGCTGGCTATGGCTCTCTTG  
AGCAGGAGCAGGCCAATAGCACTTAGCCCCGCTGGGGGCCCTAACCTCATTACCTTCTCTTGTCTGCCTCAGCC  
CCAGGAAGGGCAAGGCAAGATGGTGGACAGATAGAGAATTGTTGCTGTATTTTTTAAATATGAAAATGTTATTAA  
ACATGTCTTCTGCC

## **FIGURE 118**

MRLSSLLALLRPALPLILGLSLGCSLSLLRVSWIQGEGEDPCVEAVGERGGPQNPD SRARLD  
QSD E D F K P R I V P Y Y R D P N K P Y K K V L R T R Y I Q T E L G S R E R L L V A V L T S R A T L S T L A V A V N R T V  
A H H F P R L L Y F T G Q R G A R A P A G M Q V V S H G D E R P A W L M S E T L R H L H T H F G A D Y D W F F I M Q D D T Y  
V Q A P R L A A L A G H L S I N Q D L Y L G R A E E F I G A G E Q A R Y C H G G F G Y L L S R S L L L R L R P H L D G C R G  
D I L S A R P D E W L G R C L I D S L G V G C V S Q H Q G Q Y R S F E L A K N R D P E K E G S S A F L S A F A V H P V S E  
G T L M Y R L H K R F S A L E L E R A Y S E I E Q L Q A Q I R N L T V L T P E G E A G L S W P V G L P A P F T P H S R F E V  
L G W D Y F T E Q H T F S C A D G A P K C P L Q G A S R A D V G D A L E T A L E Q L N R R Y Q P R L R F Q K Q R L L N G Y R  
R F D P A R G M E Y T L D L L L E C V T Q R G H R R A L A R R V S L L R P L S R V E I L P M P Y V T E A T R V Q L V L P L L  
V A E A A A P A F L E A F A A N V L E P R E H A L L T L L L V Y G P R E G G R G A P D P F L G V K A A A A E L E R R Y P G  
T R L A W L A V R A E A P S Q V R L M D V V S K K H P V D T L F F L T T V W T R P G P E V L N R C R M N A I S G W Q A F F P  
V H F Q E F N P A L S P Q R S P P G P P G A G P D P P S P P G A D P S R G A P I G G R F D R Q A S A E G C F Y N A D Y L A A  
R A R L A G E L A G Q E E E E A L E G L E V M D V F L R F S G L H L F R A V E P G L V Q K F S L R D C S P R L S E E L Y H R  
C R L S N L E G L G G R A Q L A M A L F E Q E Q A N S T

### **Signal sequence:**

amino acids 1-15

### **Transmembrane domain:**

amino acids 489-507

### **N-glycosylation site.**

amino acids 121-125, 342-346

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 319-323, 464-468

### **Casein kinase II phosphorylation site.**

amino acids 64-68, 150-154, 322-326, 331-337, 368-372, 385-389,  
399-403, 409-413, 473-477, 729-733, 748-752

### **Tyrosine kinase phosphorylation site.**

amino acids 736-743

### **N-myristoylation site.**

amino acids 19-25, 23-29, 136-142, 397-403, 441-447, 544-550,  
558-564, 651-657, 657-663, 672-678

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 14-25

### **Cell attachment sequence.**

amino acids 247-250

## FIGURE 119

CGGAGTGGTGCGCCAACGTGAGAGGAAACCCGTGCGCGGCTGCGCTTTCCTGTCCCCAAGCC  
GTTCTAGACGCGGGAAAAATGCTTTCTGAAAGCAGCTCCTTTTTGAAGGGTGTGATGCTTGG  
AAGCATTTTTCTGTGCTTTGATCACTATGCTAGGACACATTAGGATTGGTCATGGAAATAGAA  
TGCACCACCATGAGCATCATCACCTACAAGCTCCTAACAAGAAGATATCTTGAAAATTTCA  
GAGGATGAGCGCATGGAGCTCAGTAAGAGCTTTCGAGTATACTGTATTATCCTTGTAACC  
CAAAGATGTGAGTCTTTGGGCTGCAGTAAAGGAGACTTGGACCAAACACTGTGACAAAGCAG  
AGTTCTTCAGTTCTGAAAATGTTAAAGTGTTTGAGTCAATTAATATGGACACAAATGACATG  
TGGTTAATGATGAGAAAAGCTTACAAATACGCCTTTGATAAGTATAGAGACCAATACAACCTG  
GTTCTTCCTTGACGCCCCACTACGTTTGCTATCATTGAAAACCTAAAGTATTTTTTGTAA  
AAAAGGATCCATCACAGCCTTCTATCTAGGCCACACTATAAAATCTGGAGACCTTGAATAT  
GTGGGTATGGAAGGAGGAATTGTCTTAAGTGTAGAATCAATGAAAAGACTTAACAGCCTTCT  
CAATATCCCAGAAAAGTGTCCTGAACAGGGAGGGATGATTTGGAAGATATCTGAAGATAAAC  
AGCTAGCAGTTTGCCTGAAATATGCTGGAGTATTTGCAGAAAATGCAGAAGATGCTGATGGA  
AAAGATGTATTTAATACCAAATCTGTTGGGCTTTCTATTAAAGAGGCAATGACTTATCACCC  
CAACCAGGTAGTAGAAGGCTGTTGTTGAGATATGGCTGTTACTTTTAATGGACTGACTCCAA  
ATCAGATGCATGTGATGATGTATGGGGTATACCGCCTTAGGGCATTTGGGCATATTTTCAAT  
GATGCATTGGTTTTCTTACCTCCAAATGGTTCTGACAATGACTGAGAAGTGGTAGAAAAGCG  
TGAATATGATCTTTGTATAGGACGTGTGTTGTCATTATTTGTAGTAGTAACATACATATCCAA  
TACAGCTGTATGTTTCTTTTTCTTTTCTAATTTGGTGGCACTGGTATAACACACATTAAAG  
TCAGTAGTACATTTTAAATGAGGGTGGTTTTTTCTTTTAAACACATGAACATTGTAAATG  
TGTTGGAAAGAAGTGTTTTAAGAATAATAATTTTGCAAATAAACTATTAATAAATATTATAT  
GTGATAAATTCTAAATTATGAACATTAGAAATCTGTGGGGCACATATTTTTGCTGATTGGTT  
AAAAAATTTTAACAGGTCTTTAGCGTTCTAAGATATGCAAATGATATCTCTAGTTGTGAATT  
TGTGATTAAAGTAAACCTTTTAGCTGTGTGTTCCCTTTACTTCTAATACTGATTTATGTTCT  
AAGCCTCCCCAAGTTCCAATGGATTTGCCTTCTCAAAATGTACAACCTAAGCAACTAAAGAAA  
ATTAAAGTGAAAGTTGAAAAAT

## **FIGURE 120**

MLSESSSFLKGVMLGSIFCALITMLGHIRIGHGNRMHHHEHHHLQAPNKEDILKISEDERME  
LSKSFRVYCIILVKPKDVSLWAAVKETWTKHCDKAEFFSSENVKVFESINMDTNDMWLMMRK  
AYKYAFDKYRDQYNWFFLARPTTFAIENLKYFLLKKDPSQPFYLGHTIKSGDLEYVGMEGG  
IVLSVESMKRLNSLLNIPEKCPEQGGMIWKISEDKQLAVCLKYAGVFAENAEDADGKDVFNT  
KSVGLSIKEAMTYHPNQVVEGCCSDMAVTFNGLTPNQMHVMMYGVYRLRAFGHIFNDALVFL  
PPNGSDND

### **Signal sequence:**

amino acids 1-33

### **N-glycosylation site.**

amino acids 121-125, 342-346

### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 319-323, 464-468

### **Casein kinase II phosphorylation site.**

amino acids 64-132, 150-154, 322-326, 331-335, 368-372, 385-389,  
399-403, 409-413, 473-477, 729-733, 748-752

### **Tyrosine kinase phosphorylation site.**

amino acids 736-743

### **N-myristoylation site.**

amino acids 19-25, 23-29, 136-142, 397-403, 441-447, 544-550,  
558-564, 651-657, 657-663, 672-672

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 14-25

### **Cell attachment sequence.**

amino acids 247-250

biochemie

## FIGURE 121

CCCACGCGTCCGATCTTACCAACAAAACACTCCTGAGGAGAAAGAAAGAGAGGGAGGGAGAG  
AAAAAGAGAGAGAGAGAGAAACAAAAAACCAAAGAGAGAGAGAAAAATGAATTCATCTAAATCAT  
CTGAAACACAATGCACAGAGAGAGGATGCTTCTCTTCCCAAATGTTCTTATGGACTGTTGCT  
GGGATCCCCATCCTATTTCTCAGTGCCTGTTTCATCACCAGATGTGTTGTGACATTTTCGCAT  
CTTTCAAACCTGTGATGAGAAAAAGTTTCAGCTACCTGAGAATTTACAGAGCTCTCCTGCT  
ACAATTATGGATCAGGTTCAAGTCAAGAAATGTTGTCCATTGAACTGGGAATATTTTCAATCC  
AGCTGCTACTTCTTTTCTACTGACACCATTTCTGGGCGTTAAGTTTAAAGAACTGCTCAGC  
CATGGGGGCTCACCTGGTGGTTATCAACTCACAGGAGGAGCAGGAATTCCTTTCTTACAAGA  
AACCTAAAATGAGAGAGTTTTTTTATTGGACTGTCAGACCAGGTGTGCGAGGGTCAGTGGCAA  
TGGGTGGACGGCACACCTTTGACAAAGTCTCTGAGCTTCTGGGATGTAGGGGAGCCCAACAA  
CATAGCTACCCTGGAGGACTGTGCCACCATGAGAGACTCTTCAAACCCAAGGCAAAATTGGA  
ATGATGTAACCTGTTTCCTCAATTATTTTCGGATTTGTGAAATGGTAGGAATAAATCCTTTG  
AACAAAGGAAAATCTCTTTAAAGAACAGAAGGCACAACCTCAAATGTGTAAAGAAGGAAGAGCA  
AGAACATGGCCACACCCACCGCCCCACACGAGAAATTTGTGCGCTGAACTTCAAAGGACTTC  
ATAAGTATTTGTTACTCTGATACAAATAAAAAATAAGTAGTTTTAAATGTTAAAAAAAAAAAAA  
AAA  
AAAAA

the "old" 500

## FIGURE 122

MNSSKSSETQCTERGCFSSQMFLWTVAGIPILFLSACFITRCVVTFRIFQTCDEKKFQLPEN  
FTELSYNYGSGSVKNCCPLNWEYFQSSCYFFSTDTISWALSLKNCSAMGAHLVVINSQEEQ  
EFLSYKKPKMREFFIGLSDQVVEGQWQWVDGTPLTKSLSFWDVGEPNNIATLEDCATMRDSS  
NPRQNWNDVTCFLNYFRICEMVGINPLNKGKSL

### **Signal sequence:**

amino acids 1-42

### **N-glycosylation site.**

amino acids 2-6, 62-66, 107-111

### **Casein kinase II phosphorylation site.**

amino acids 51-55, 120-124, 163-167, 175-179, 181-185

### **N-myristoylation site.**

amino acids 15-21, 74-80, 155-161

### **Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 27-38

Protein Data Bank

# FIGURE 123

GGGACTACAAGCCGCGCCGCGCTGCCGCTGGCCCCCTCAGCAACCCTCGACATGGCGCTGAGGCGGCCACCGCGAC  
TCCGGCTCTGCGCTCGGCTGCCTGACTTCTTCCTGCTGCTGCTTTTCAGGGGCTGCCTGATAGGGGCTGTAAATC  
TCAAATCCAGCAATCGAACCCAGTGGTACAGGAATTTGAAAGTGTGGAAGTGTCTTGATCATTACGGATTTCGC  
AGACAAAGTGACCCAGGATCGAGTGGAAGAAAATTCAGATGAACAAACCATATGTGTTTTTTGACAACAAAA  
TTCAGGGAGACTTGGCGGGTCTGTGCAGAAATACTGGGGAAGACATCCCTGAAGATCTGGAATGTGACACGGAGAG  
ACTCAGCCCTTTATCGCTGTGAGGTGCTTGCTCGAAATGACCGCAAGGAAATTGATGAGATTGTGATCGAGTTAA  
CTGTGCAAGTGAAGCCAGTGACCCCTGTCTGTAGAGTGCCGAAGGCTGTACCAGTAGGCAAGATGGCAACACTGC  
ACTGCCAGGAGAGTGAGGGCCACCCCCGGCCTCACTACAGCTGGTATCGCAATGATGTACCACTGCCACGGATT  
CCAGAGCCAATCCCAGATTTTCGCAATTCTTCTTTCCACTTAAACTCTGAAACAGGCACCTTTGGTGTTCACTGCTG  
TTCACAAGGACGACTCTGGGCAGTACTACTGCATTGCTTCCAATGACGCAGGCTCAGCCAGGTGTGAGGAGCAGG  
AGATGGAAGTCTATGACCTGAACATTGGCGGAATTATTGGGGGGGTTCTGGTTGTCTTGTCTGTAAGTGGCCCTGA  
TCACGTTGGGCATCTGCTGTGCATACAGACGTGGCTACTTTCATCAACAATAAACAGGATGGAGAAAGTTACAAGA  
ACCCAGGGGAAACCAGATGGAGTTAACTACATCCGCACTGACGAGGAGGGCGACTTCAGACACAAGTCACTCGTTTG  
TGATCTGAGACCCGCGGTGTGGCTGAGAGCGCACAGAGCGCACGTGCACATACCTCTGCTAGAAAACCTCTGTCAA  
GGCAGCGAGAGCTGATGCACTCGGACAGAGCTAGACACTCATTTCAGAAGCTTTTCGTTTTGGCCAAAGTTGACCA  
CTACTCTTCTTACTCTAACAAGCCACATGAATAGAAGAATTTTCCTCAAGATGGACCCGGTAAATATAACCACAA  
GGAAGCGAAACTGGGTGCGTTCACTGAGTTGGGTTCCCTAATCTGTTTTCTGGCCTGATTCCCGCATGAGTATTAGG  
GTGATCTTAAAGAGTTTGCTCAGTAAACGCCCGTGCTGGGCCCTGTGAAGCCAGCATGTTCAACCACTGGTCTGTT  
CAGCAGCCACGACAGCACCATTGTGAGATGGCAGGTGGCTGGACAGCACAGCAGCGCATCCCGCGGGAACCCA  
GAAAAGGCTTCTTACACAGCAGCCTTACTTTCATCGGCCACAGACACCACCGCAGTTTCTTTTAAAGGCTCTGC  
TGATCGGTGTTGCAGTGTCCATTGTGGAGAAGCTTTTGGATCAGCATTTTGTAAAAACAACCAAAATCAGGAAG  
GTAAATTTGGTTGCTGGAAGAGGGATCTTGCTGAGGAACCCTGCTTGTCCAACAGGGTGTGAGGATTTAAGGAAA  
ACCTTCGCTCTTAGGCTAAGTCTGAAATGGTACTGAAATATGCTTTTCTATGGGTCTTGTATTATTTTATAAAATTT  
TACATCTAAATTTTGTCTAAGGATGTATTTTGTATTATGAAAAGAAAATTTCTATTAAACTGTAAATATATTGT  
CATACAATGTTAAATAACCTATTTTTTTTAAAAAAGTTCACTTAAAGGTAGAAGTTCCAAGTACTAGTGTAAAT  
TGGAAAAATATCAATAATTAAGAGTATTTTACCCAAGGAATCCTCTCATGGAAGTTTACTGTGATGTTCTTTTCT  
CACACAAGTTTTAGCCTTTTTTACAAGGGAACCTACTGTCTACACATCAGACCATAGTTGCTTAGGAAACCTT  
TAAAAATTCAGTTAAGCAATGTTGAAATCAGTTTGCTCTCTTCAAAAGAAACCTCTCAGGTAGCTTTTGAAC  
GCCTCTTCTGAGATGACTAGGACAGTCTGTACCCAGAGGCCACCCAGAAGCCCTCAGATGTACATACAGATG  
CCAGTCAGCTCCTGGGGTTGCGCCAGGCGCCCCCGCTCTAGCTCACTGTTGCCTCGCTGTCTGCCAGGAGGCCCT  
GCCATCCTTGGGCCCTGGCAGTGGCTGTGTCCAGTGAGCTTACTCACGTGGCCCTTGCTTCATCCAGCACAGC  
TCTCAGGTGGGCACTGCAGGGACACTGGTGTCTTCCATGTAGCGTCCAGCTTTGGGCTCCTGTAAACAGACCTCT  
TTTTGGTTATGGATGGCTCACAAAATAGGGCCCCCAATGCTATTTTTTTTTTTTAAAGTTTGTTTAATTATTTGTT  
AAGATTGTCTAAGGCCAAAGGCAATTGCGAAATCAAGTCTGTCAAGTACAATAACATTTTTTAAAGAAAATGGAT  
CCCACTGTTCTCTTTGGCCACAGAGAAAGCACCCAGACGCCACAGGCTCTGTGCAATTTCAAAACAAACCATGAT  
GGAGTGGCGGCCAGTCCAGCCTTTTAAAGAACGTGAGTGGAGCAGCCAGGTGAAAGGCCTGGCGGGGAGGAAAG  
TGAAACGCCTGAATCAAAAGCAGTTTTTCTAATTTTGACTTTAAATTTTTTTCATCCGCCGAGACACTGCTCCCAT  
TGTGGGGGGACATTAGCAACATCACTCAGAAGCCTGTGTTCTTCAAGAGCAGGTGTTCTCAGCCTCACATGCCCT  
GCCGTGCTGGACTCAGGACTGAAGTGTCTGTAAAGCAAGGAGCTGCTGAGAAGGAGCACTCCACTGTGTGCCCTGGA  
GAATGGCTCTCACTACTCACCTTGTCTTTCAGCTTCCAGTGTCTTGGGTTTTTTTATACTTTGACAGCTTTTTTTT  
AATTGCATACATGAGACTGTGTTGACTTTTTTTTAGTTATGTGAAACACTTTGCGCGAGGCCGCTGGCAGAGGCA  
GGAAATGCTCCAGCAGTGGCTCAGTGTCTCCCTGGTGTCTGCTGCATGGCATCCTGGATGCTTAGCATGCAAGTTC  
CCTCCATCATTGGCACCTTGGTAGAGAGGGATGGCTCCCCACCTCAGCGTTGGGGATTACGCTCCAGCCTCCT  
TCTTGGTTGTATAGTATAGGGTAGCCTTATTGCCCCCTCTTCTTATACCCTAAAACCTTCTACACTAGTGCCA  
TGGGAACCAAGTCTGAAAAAGTAGAGAGAAGTGAAGTAGAGTCTGGGAAGTAGCTGCCTATAACTGAGACTAGA  
CGGAAAAGGAATACTCGTGTATTTTAAAGATATGAATGTGACTCAAGACTCGAGGCCGATACGAGGCTGTGATTCT  
GCCTTTGGATGGATGTTGCTGTACACAGATGCTACAGACTTGTACTAACACACCGTAATTTGGCATTTGTTTAAAC  
CTCATTATATAAAGCTTCAAAAAAACCA

## **FIGURE 124**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77624

><subunit 1 of 1, 310 aa, 1 stop

><MW: 35020, pI: 7.90, NX(S/T): 3

MALRRPPRLRLCARLPDFFLLLLFRGCLIGAVNLKSSNRTPVVQEFESVELSCIITDSQTS  
PRIEWKKIQDEQTTYVFFDNKIQGDLAGRAEILGKTSCLKIWNVTRRDSALYRCEVVARNDRK  
EIDEIVIELTVQVKPVPVCRVPKAVPVGKMATLHCQESEGHPRPHYSWYRNDVPLPTDSRA  
NPRFRNSSFHLNSETGTLVFTAVHKDDSGQYYCIASNDAGSARCEEQEMEVDLNIIGGIIGG  
VLVVLAVLALITLGICCAYYRRGYFINNKQDGESYKNPGKPDGVNYIRTDEEGDFRHKSSFVI

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-30

#### **Transmembrane domain:**

amino acids 243-263

#### **N-glycosylation sites.**

amino acids 104-107, 192-195

#### **cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 107-110

#### **Casein kinase II phosphorylation site.**

amino acids 106-109, 296-299

#### **Tyrosine kinase phosphorylation site.**

amino acids 69-77

#### **N-myristoylation sites.**

amino acids 26-31, 215-220, 226-231, 243-248, 244-249, 262-267

biochem 2000